Countering Misinformation Early: Evidence from a Classroom-Based Field Experiment in India*

Priyadarshi Amar† Sumitra Badrinathan‡ Simon Chauchard§ Florian Sichart¶

Abstract

As evidenced during covid-19, medical misinformation can have severe consequences, especially among populations with low levels of digital literacy and education. This paper evaluates whether a sustained, classroom-based media literacy education can help school children in Bihar, India, become more discerning consumers of misinformation. Collaborating with Bihar's state government, we tested a field experiment in 583 villages, targeting over 12,000 students, that engaged treatment respondents in classroom discussions about health misinformation topics. Results from intent-to-treat estimates demonstrate that our intervention had significant impacts on a number of counts: treated students showed heightened ability to discern between true and false information, changed health preferences, increased reliance on scientific medicine and lower dependence on untrustworthy sources of news. We resurveyed participants 4 months after the program and found that not only did effects persist, treated respondents were able to discern true from false political information as well. Finally, we also detect spillover network effects within families: parents of treated students are significantly better at discerning true from false information. These findings hold significant implications for assessing the efficacy of media literacy measures at scale in light of previous studies that often yield null or mixed results. Our paper speaks to the ability of policy-based interventions to have externally valid and long-term results.

Keywords: Misinformation, Field Experiment, Education, India

*Authors listed alphabetically. This study is funded by the Social Science Research Council. We are grateful to Jeevika and especially to IAS officer Mr. Rahul Kumar, Pushpendra Singh Tiwari, and Rakesh Kumar for facilitating the Bihar government's official collaboration on this project. Our teacher training and implementation partner is DataLeads in New Delhi, India. Our survey partner is Sunai Consultancy in Patna, India. This research received IRB approval from American University and the UC3M's data protection officer. We thank Josiah Gottfried for outstanding research assistance, and Minati Chaklanavis, Pranav Chaudhary and Nishant Attri for on-field support. Pre-registration available at OSF: https://osf.io/h43qn. For comments and feedback we are grateful to Laura Paler, Austin Davis, Gareth Nellis, Erik Kramon, Laura Schechter, Horacio Larreguy, Cyrus Samii, Casey Petroff, Maria Hernandez de Benito and seminar participants at George Washington, Penn, American, Rochester, IC3JM, CUNEF MAPE, the India-China conference (Namur), APSA, and UMontreal.

†PhD Candidate, University of Wisconsin-Madison

‡Assistant Professor, American University

§Associate Professor, University Carlos III Madrid & Instituto Carlos 3 - Juan March

¶PhD Candidate, Princeton University

1 Introduction

The dissemination of misinformation has garnered considerable attention on a global scale (Persily, Tucker, and Tucker, 2020). However, misinformation presents a particularly acute and pressing challenge within developing and emerging economies. These contexts are characterized by lower levels of digital literacy, burgeoning internet access, informal information networks, and the instrumentalization of misinformation along identity lines (Badrinathan, 2021; Chauchard and Garimella, 2022; Gottlieb, Adida, and Moussa, 2022). Despite significant scholarly advancements in the realm of misinformation mitigation strategies, the predominant focus of this literature remains on Western contexts, resulting in a paucity of understanding regarding effective approaches for combating misinformation within the Global South (Blair et al., 2023; Badrinathan and Chauchard, 2023a). India, the focus of this study, exemplifies the multifaceted nature of this problem. Misinformation in key areas such as politics and health can have severe downstream effects, influencing political violence and health-seeking behaviors, including vaccination and disease treatment (Bridgman et al., 2020). Low state capacity, especially in healthcare, exacerbates the problem, while diminishing spaces for independent media further weakens the flow of credible information (Mohan, 2021; Sen, 2023; Sharma, 2015). Compounding these issues is the role of political and religious elites, who often generate and spread misinformation to serve partisan agendas, making it more likely to be believed due to its elite origins (Siddiqui, 2020). This confluence of elite-backed disinformation, limited access to credible news, and institutional weaknesses creates an environment where misinformation proliferates unchecked.

How can we address misinformation in such a context? To answer this question, we implemented a sustained, classroom-based education field experiment in Bihar, one of India's least developed states, targeting over 12,000 adolescents aged 13–18 across 583 villages.

A substantial body of literature has explored various interventions to mitigate misinformation. These include fact-checking initiatives (Clayton et al., 2019), providing corrections (Porter and Wood, 2019; Bode and Vraga, 2018), emphasizing information accuracy (Pennycook and Rand, 2019), and preemptive debunking (Roozenbeek et al., 2022), among other approaches. Yet, while some of these interventions have demonstrated efficacy, they often constitute one-off, online interventions tailored to digitally literate urban samples and rarely adapted to the

most vulnerable populations (Brashier, 2024) or to populations exposed to misinformation offline. Moreover, research on interventions within the Global South is notably scant, with findings generally yielding mixed or negligible effects (Blair et al., 2023). Outside of academic scholarship, governments, NGOs, and civil society entities have increasingly championed iterative, classroom-based media and information literacy initiatives, most often among children and teenagers, both in the Global North and South. For instance, legislative efforts in New Jersey seek to mandate media literacy education for K-12 students (Sitrin, 2020), mirroring similar initiatives across Africa and South Asia (Blair et al., 2023). Crucially, however, empirical evidence estimating the causal impact of such classroom-based media literacy training among school children is so far almost entirely lacking.[1]

In this study, we attempt to address this gap by exploring this less conventional and understudied approach to counter misinformation, namely, sustained education on media and information literacy within classroom settings. Our intervention, which we evaluate through a randomized controlled trial, was deployed across 583 villages throughout Bihar in India, targeting over 12,000 children. Aimed at students in grades 8 through 12, the intervention facilitated classroom-based sessions addressing misinformation. These classes, conducted over a period of 14 weeks, entailed four 90-minutes sessions held roughly every three weeks, as well as homework assignments between these sessions. The curriculum, which we custom-designed for this study, focused substantively on health misinformation and highlighted imparting factual knowledge and endorsing scientific beliefs while educating students on the factors contributing to susceptibility to misinformation, its ramifications, and strategies for discerning dubious information. The goals of the curriculum were to (1) impart new knowledge about misinformation, (2) provide skills and tools that could be applied to non-classroom settings, and (3) change norms around misinformation.

Our intervention is, by design, a bundled treatment combining several elements into a comprehensive learning course similar to other education-based experiments (Dhar, Jain, and

[1]We note a single exception: Apuke, Omar, and Asude Tunca (2023) conduct a media literacy intervention substantively comparable to ours by enrolling 470 students at a single public university in Nigeria in a study where treatment group respondents received 6 weeks of classes. While the training found beneficial effects, we note the limited sample size and scope of the study, as well as lack of pre-registration of the experiment and discussion on internal validity issues such as spillover effects or compliance with treatment.

Jayachandran, 2022; Badrinathan, 2021). We recruited and trained special purpose educators to deliver lessons in government-owned libraries across the state. The Bihar Rural Livelihoods Promotion Society (BRLPS, or as it is commonly known, Jeevika), an autonomous body under the Bihar State Government, sanctioned the program to be administered as an official government certificate course, thereby extending the reach and legitimacy of the intervention, and boosting enrollment. We randomized respondents into treatment at the village level, with the study targeting one classroom size (20 to 24 respondents) per village. Control group villages received placebo classes on very basic conversational English, designed to hold constant attendance and compliance with a long-term program and only vary the content of instruction.

We evaluate the effect of this intervention on a number of misinformation outcomes, including attitudes, preferences, and behaviors. Intent-to-treat estimates measured soon after the intervention indicate that the intervention had a strong, significant, impact on students' capacity to comprehend and process information, as well as to apply classroom teachings to real-life contexts. At the conclusion of the curriculum, treated participants demonstrated heightened discernment in evaluating information and making decisions regarding the sharing of news items (0.32 SD), with effects sizes substantially larger than those previously identified. Notably, the intervention also brought about changes in their health preferences (0.21 SD), diminishing reliance on alternative medical approaches to cure serious illnesses. This is crucial as previous studies show that misinformation related to health, and especially misinformation tied to traditional beliefs systems, is more difficult to change (Chauchard and Badrinathan, 2024). Further, the treatment enhanced assessment of sources and the ability of treated students to discern their credibility. Finally, while the treatment did not change behaviors around engaging with misinformation countermeasures on average, it did so among boys, demonstrating that behavioral change in such contexts may be more challenging when conservative gender norms serve as barriers.

What is especially striking is that these effects persisted over time. We resurveyed a random sub-sample of 2,000 participants about 4 months after the intervention and continued to find a large effect on ability to discern true from false information (0.27 SD). Crucially, our second endline included a battery of political news headlines that were not discussed in the classroom and not included in the first endline. We find that there are large spillover effects on these entirely new items – respondents are able to discern true from false political news 6 months after

an intervention that focused on health misinformation (0.31 SD), demonstrating that they were able to learn from the treatment, retain its lessons, and apply it to entirely new, and polarizing, domains. Finally, we also find that parents of treated students are better able to discern true from false information, demonstrating the ability of sustained educative interventions to have network-level spillover effects.

The implications of these findings hold significant importance, especially in light of prevailing challenges faced by most media or information literacy interventions, which often yield null or mixed results (Blair et al., 2023). Our intervention demonstrates effectiveness among a particularly challenging sample and in a notably arduous context. Bihar typically exhibits low prioritization of education, contributing to a substantial dropout rate in public schools, especially among girls (Muralidharan and Prakash, 2017). Moreover, standard of living indicators for Bihar consistently underscore its under-performance compared to the rest of the country in various metrics, primarily attributable to weak state capacity (Desai, 2010). Within this context of state dysfunction, it is notable that public programs across the board demonstrate weak performance (Rasul and Sharma, 2014; Mathew and Moore, 2011; Jha, 2023). Thus, it was not readily foreseeable that our intervention would yield such significant outcomes. Despite the challenging context and the escalating politicization of health issues across India, adolescents demonstrated engagement with the class, which in turn influenced their preferences and behavioral intentions.

Our study contributes to the expanding body of literature addressing strategies to mitigate misinformation on a global scale. But what distinguishes our research from existing work is its departure from the assumption that misinformation predominantly manifests as an online phenomenon. Prior research has largely operated under this presumption, leading to the development of numerous interventions tailored to online demographics. These interventions include, for instance, online gaming as a strategy to combat misinformation (Roozenbeek and van der Linden, 2019) and the provision of corrective nudges through direct messaging on Twitter (Pennycook et al., 2021). However, such interventions are inherently ill-suited for individuals who do not have regular or sustained access to the internet. Moreover, in contexts like India, those who do have access often use encrypted messaging platforms like WhatsApp, where interventions reliant on platform-based changes are not feasible (Badrinathan, 2021). Indeed, prior research in developing contexts has underscored the significance of offline sources and commu-

nal spaces for news dissemination and consumption, such as interpersonal communication and face-to-face interactions in community settings like marketplaces (Gadjanova, Lynch, and Saibu, 2022). Operating under this premise, one of our primary contributions is studying adolescents. This approach enables us to target a demographic with potentially less entrenched attitudes, significant because studies on misinformation consistently indicate that a key impediment to the effectiveness of interventions is motivated reasoning, the inclination to accept or reject information based on pre-existing beliefs and identities (Taber and Lodge, 2006). This phenomenon is especially pronounced among individuals who exhibit political sophistication, possess strong political affiliations, and harbor extensive political knowledge (Li and Wagner, 2020). By focusing on adolescents, we potentially engage with a population that has not yet developed resistance to altering entrenched viewpoints (Niemi and Jennings, 1991).

Finally, collaborating directly with a government agency to roll out this RCT as an official government program increased the legitimacy and reach of our treatment, helping dilute non-compliance, and ultimately helped resemble an "in-the-wild" roll out of a school program. This study, thus, has significant broader implications not just for countering misinformation but also, given its persistent effects, for the creation of education policy and public health strategies in developing countries.

2 Sustained Education Against Misinformation

The global rise of misinformation has attracted substantial academic and policy attention (Persily, Tucker, and Tucker, 2020), sparking a surge of experimental studies aimed at identifying solutions. However, many of these interventions have limitations. First, the effects of most interventions do not succeed in persisting over time (Porter and Wood, 2023). Second, they predominantly target digital and online populations, based on the assumption that misinformation is an online-only phenomenon. This overlooks significant offline populations, especially in rural or underconnected areas (Gadjanova, Lynch, and Saibu, 2022). In many developing countries even when personalized access to the internet exists, this takes place largely through the medium of WhatsApp, rather than open platforms such as Twitter or Facebook. WhatsApp, characterized by encrypted peer to peer sharing, makes online interventions such as such as cor-

rections and fact-checking, or algorithmic changes, inappropriate and unfeasible (Badrinathan, 2021; Rossini et al., 2020; Valeriani and Vaccari, 2018).[2] Moreover, most interventions rely on one-time nudges or priming techniques that, while useful for short-term behavior modification, fail to address deeper, entrenched norms or build lasting skills (for example see Pennycook and Rand (2019)). Even among interventions aimed at more sustained change, such as inoculation strategies (Roozenbeek et al., 2022), success has been elusive in the Global South. The bulk of this research remains concentrated in the Global North, with mixed results when tested in non-Western settings, particularly in developing economies (Blair et al., 2023; Badrinathan and Chauchard, 2023a).

Yet, misinformation poses a particularly urgent threat in low and middle income countries. These countries often exhibit lower levels of digital literacy, rapidly expanding internet access, and a reliance on informal information networks. Misinformation in these contexts can be further weaponized along identity lines, exacerbating social and political divisions (Badrinathan, 2021; Chauchard and Garimella, 2022; Gottlieb, Adida, and Moussa, 2022). To address these challenges, we designed a classroom-based field experiment, relying on a program entitled the Bihar Information and Media Literacy Initiative (BIMLI). The intervention had two key goals: (a) increasing knowledge via skills and facts, and (b) shifting norms related to misinformation.

By *knowledge*, we refer to two key components: (1) recall—the ability to remember specific information learned in the classroom, and (2) application—the ability to use skills and tools acquired in class to critically assess new information beyond the classroom. In the first case, the intervention may enhance knowledge by exposing students to specific, accurate facts they can later recall. For example, since the curriculum focuses on health-related misinformation, students might remember that there is no evidence supporting the claim that papaya leaves cure dengue fever. In the second case, students may develop more general knowledge, such as effective strategies for detecting misinformation in the future or consuming different types of

[2]While scholars have looked at social corrections or peer corrections to misinformation on encrypted apps (Badrinathan and Chauchard, 2023b; Bode and Vraga, 2015), such techniques assume that people already have the knowledge, desire and skills to fact check information, such that when they receive primes about accuracy or corrections, they are able to change beliefs and behavior. These assumptions may not adequately apply in the case of populations that are new to the internet, have low levels of digital literacy, and reside in rural areas with low socio-economic backgrounds.

information critically. These broader skills could significantly influence how they evaluate new information they encounter even after the program.

With respect to *norms*, we designed the intervention to influence behaviors and attitudes by shaping perceptions of misinformation, drawing on the idea that institutions can serve as powerful sources of normative influence (Tankard and Paluck, 2016). Similar to classroom-based interventions on gender norms by Dhar, Jain, and Jayachandran (2022), our program's affiliation with a state agency and the involvement of local authority figures likely reinforced signals about appropriate attitudes and behaviors. Additionally, teachers—often local role models—may have been seen as influential social referents for students (Paluck and Shepherd, 2012). Several elements in the curriculum explicitly or implicitly targeted normative change, including modules designed to (1) elevate the perceived importance of misinformation as a societal issue, (2) increase awareness of the risks associated with misinformation, and (3) shift what individuals consider acceptable to say, believe, and share within their social networks.

Overall, the BIMLI program was thus designed to achieve multiple objectives. It sought not only to enhance knowledge — both specific and general — about misinformation but also to shift social norms around how misinformation is perceived and responded to. By targeting both cognitive and normative dimensions, our goal was to foster long-lasting changes in behaviors and attitudes, addressing the deeper, more ingrained habits that contribute to the persistence of misinformed beliefs.

The BIMLI program was thus, by design, a significantly more intensive and substantively different intervention than those typically tested in the existing literature. To ensure it succeeded while addressing limitations of previous work, we introduced several key innovations. First, the mode of delivery. We administered the program face-to-face in classroom settings, fostering a peer-based, interactive environment with iterative learning, where respondents encountered key lessons repeatedly over multiple sessions. Research suggests that peer interactions in classroom settings can deepen understanding by exposing learners to diverse perspectives and fostering critical discussion (Dhar, Jain, and Jayachandran, 2022), while repeated exposure allows for reinforcement and application of concepts in varied contexts (Fazio, Rand, and Pennycook, 2019). Additionally, we partnered with an arm of the Bihar state government to roll out the program as an official government-endorsed certificate course, thereby increasing its legitimacy. This official

recognition incentivized participation and ensured higher compliance among respondents and facilitators alike.

Our second innovation was with regards to the intensity of the treatment. The total dosage of our program amounted to approximately 10 contact hours between facilitators and students. This is significantly more labor-intensive than many existing studies testing education and information-based interventions. For example, Hameleers (2020) and Guess et al. (2020) provided respondents with tips to spot misinformation via flyers, while Ali and Qazi (2021) used a brief 3-4 minute media literacy video. Closer in design to our study, Badrinathan (2021) engaged respondents in a one-hour discussion on media literacy, and Moore and Hancock (2022) offered a self-directed hour-long digital literacy module. However, these are typically short-term, one-off treatments without sustained engagement, and many other interventions rely on primes or reminders about critical thinking rather than providing actual tools or techniques for long-term learning. Consequently, we opted for this sustained, iterative approach involving multiple sessions of learning as well as homework assignments between sessions, thereby departing from almost all other educative interventions in the misinformation literature (Blair et al., 2023), aimed at creating substantial variation in the independent variable (Mutz, Druckman, and Green, 2021).

Third, our target population was adolescents. Focusing on adolescents is crucial for several reasons. Existing research shows that adults, especially in polarized contexts, are significantly harder to influence due to cognitive biases such as motivated reasoning (Badrinathan, 2021). Adolescents, by contrast, are still forming their worldviews and may be more open to new information and perspectives. This demographic, therefore, offers a unique opportunity to intervene before attitudes and behaviors related to misinformation become deeply entrenched (Niemi and Jennings, 1991).

3 The Politics of Misinformation in India

Health-related misinformation is widespread in India. From our own control group data, 87% of respondents reported believing in the curative powers of papaya leaves for dengue fever. In other studies from similar contexts (Chauchard and Badrinathan, 2024), over 60% of respondents claimed that cow urine could cure COVID-19. This type of misinformation can have severe

consequences—it discourages people from seeking actual medical solutions, such as hospital treatment, which can sometimes lead to fatal outcomes (Bridgman et al., 2020). The negative consequences of belief in misinformation may be particularly pronounced in regions with low education and literacy rates.

These deeply entrenched beliefs are also tied to polarizing identities in India, as they are often manipulated by political elites to gain support. Traditional health remedies, many of which are linked to ancient Hindu culture, are used as a tool to drum up support from the Hindu voter base. This is especially evident with the rise to power of the Hindu nationalist Bharatiya Janata Party (BJP), which has positioned itself as a champion of Hindu values (Jaffrelot, 2021). A recent example elite-driven health disinformation involves a BJP politician hosting a public event where participants drank cow urine, believing it to be a cure for COVID-19—an event that ended with several attendees being hospitalized (Siddiqui, 2020). Politicians from the BJP in this context have even boasted about their ability to make anything go viral, true or false (Kumar, 2019). Previous research demonstrates that misinformation that resonates with long-standing identities can be particularly hard to dislodge (Nyhan, 2021; Chauchard and Badrinathan, 2024). Indian politics has long been dominated by a fundamental cleavage between Hindus and Muslims, and the prominence of religion as a social identity has been central (Brass, 2005). Thus it stands to reason that misinformed beliefs, particularly when propagated by elites and seeking to divide already polarized populations further along existing cleavages, can be very difficult to change.

For citizens in such contexts, finding ways out of the misinformation trap is challenging. This is particularly true in Bihar, the site of our intervention. Bihar is India's poorest state, home to 127 million people, with over one-third of the population living below the poverty line. A child born in Bihar has a lower likelihood of surviving to age five than a child born in Afghanistan (Kumar, Piyasa, and Saikia, 2022; Rampal, 2023). The state's relative underdevelopment translates into a lack of essential services such as healthcare and education, alongside the failure of many public programs (Sharma, 2015). Children in Bihar, especially girls, are significantly less likely to attend school compared to those in other states (Muralidharan and Prakash, 2017). Access to the internet is also limited; in our baseline data, less than 20% of our sample reported having used a cellphone. Even in households with internet access, it was often through a shared mobile phone, highlighting the stark contrast with Western settings, where internet access is more

individualized (Rangaswamy and Singh, 2009). This limited connectivity is compounded by a deteriorating informational environment across India. Independent media and spaces for dissent are shrinking, as state capture of institutions, including news sources, intensifies (Mohan, 2021; Sen, 2023). This aligns with broader observations of democratic decline in the country recent years, where the space for credible information has narrowed significantly alongside decreasing state capacity (Tudor, 2023).

Thus, Bihar presents a tough test case for changing knowledge and norms around misinformation for the better. However, it is precisely in such contexts that interventions to counter misinformation can have the greatest impact. The challenge is heightened by our focus on health-related misinformation, where not only political elites but also long-standing cultural myths and generational beliefs contribute to the spread of misinformation.

4 Experimental Design and Data Collection

We implement a randomized control trial to test the efficacy of the BIMLI program in a sample of 583 villages across 32 districts of the state of Bihar. The unit of randomization into treatment was the village. While respondents in treatment villages received four sessions of media literacy classroom lessons, control respondents received four sessions of conversational English classroom lessons (described in detail below).

4.1 The Treatment

The BIMLI program featured four classroom sessions, each about 90 minutes long and approximately 2-3 weeks apart, as well as homework assignments between sessions. We created a custom curriculum and lesson plan for BIMLI for this study. In doing so, our educative curriculum, though bundled, focused broadly on media and information literacy and critical thinking, with the goal of changing norms and providing knowledge. Media and information literacy is broadly defined as the skills and competencies that promote critical and responsible engagement with information and media sources around us (Jones-Jang, Mortensen, and Liu, 2021). In practice, however, this concept is operationalized in a myriad of ways, with scholars prioritizing different elements of the concept such as critical thinking, knowledge about media industries,

or knowledge about psychological biases in thinking. In Table 1 we provide a summary of our treatment lesson plan, including a brief description of learning goals for each lesson, as well as identifying the theoretical framework on which each topic is based.

We worked with local Bihar educators and experts in the media literacy in India to develop this curriculum, including modules that translate the learning objectives for each session into actionable lesson plans for teachers. In doing so, we ensured examples were tailored to the local context along with discussion and activity-based lessons. For example, we incorporated role-playing in the classroom to underscore how talking to adults about health misinformation can be challenging, especially in the case of dislodging traditional alternative medical practices and beliefs (Malhotra and Pearce, 2022). The extensive materials we developed for this project — including detailed slide decks, guidelines for teachers dealing with children, time-use lesson plans, and photo banks with examples — are tailored to the local context (and delivered in Hindi) but we hope they can provide a base and be a resource for media literacy trainings across the world. Appendix B includes an overview of the detailed lesson plan and samples of materials from the treatment.

Apart from the four in-person sessions, to complement in-class materials and encourage further reflection we also designed homework assignments that students completed between sessions. These included writing stories and recording observations, along with talking to family members about what was learned in the classroom. Finally, after each session, we also provided succinctly designed take-home summary sheets for students that recap the goals and objectives of each lesson, designed to act as a guide or reference booklet.

4.2 Administering classes

To bolster the credibility of BIMLI, we secured official collaboration with an autonomous agency of the Bihar local government, the Bihar Rural Livelihoods Promotion Society (BRLPS, or as it is commonly known, Jeevika), to launch BIMLI statewide. Despite their governmental affiliation, Jeevika operates autonomously under the leadership of an Indian Administrative Services officer. Our engagement with Jeevika commenced in mid-2022, driven by a shared interest in combating the proliferation of misinformation, particularly the prevalent reliance on homemade remedies

Table 1: Description and Objectives of Media Literacy Curriculum

	Learning Objectives	Theoretical Framework	Indian Context
Module 1	Definitions: What misinformation is and what it isn't; where we encounter it; consequences on health, intergroup harmony, elections. Increasing awareness about the existence and dangers of misinformation	Guess and Lyons (2020)'s definition of misinformation and several examples of recent misinformation and its effects	Examples: health misinformation leading to vaccine hesitancy, falsehoods on WhatsApp groups, photoshopped images
Module 2	Causes of vulnerability: human psychological biases like confirmation bias; news and media system biases such as financial incentives; the need for critical thinking	Motivated reasoning from Taber and Lodge (2006), list of psychological biases adapted from Roozenbeek et al. (2022), list of media biases adapted from Ashley, Maksl, and Craft (2013), news framing effects from Druckman and Nelson (2003), fact-opinion discernment from Graham and Yair (2023)	Introduction to the news media environment in India + how biases manifest in the Indian context (for example, scapegoating minorities)
Module 3	Solutions to misinformation: what makes a reliable source; discerning true from false information; authority sources and their power; impulsive sharing behavior	Concrete examples of tips to spot misinformation (Guess, Nagler, and Tucker, 2019; Vraga, Bode, and Tully, 2022; Badrinathan, 2021), focus on sharing as different from belief (Brashier and Schacter, 2020)	Tailored Indian examples focused on WhatsApp such as reverse image search, looking for the 'forwarded many times' tag, introduction to fact-checking sources
Module 4	Family, vaccination: how to talk to friends and relatives who spread misinformation; the benefits of vaccinations (Covid and others); recap of all previous learning	Efficacy of social corrections (Bode and Vraga, 2018; Badrinathan and Chauchard, 2023b), talking to older family about misinformation (Pearce and Malhotra, 2022)	Role-playing exercise to talk to an Indian relative about misinformation

and alternative medicine sources among youth, and especially in rural regions.

In our study, participants were school students in grades 8 through 12, aged between 13 and 18 years old and the classes took place from November 2023 to March 2024. We hired and trained special-purpose teachers to deliver the classes, each of whom visited each classroom roughly once every two to three weeks. We partnered with DataLeads, a Delhi-based organization experienced in media literacy and fact-checking, to hire teachers for our intervention.[3]

To dispense the classes, Jeevika made available to us 99 community libraries across 32 districts in Bihar (one in each block). We ran our classes in these libraries. We opted not to use existing public school teachers or school settings for this program. Our meetings with government officials in Bihar revealed that public school teachers were already overburdened, that absenteeism (in students and teachers) was common in Bihar, and that running the intervention in schools may increase the risk of spillover.[4] Further, adding extra modules and hours of instruction to public school curriculum would have require extensive bureaucratic authorization. Recognizing that the success of such a program is contingent not only on student attendance and compliance, but also crucially on teacher attendance, we hired a separate pool of teachers for this intervention. Given the nature of (lack of) infrastructure and resources in Bihar, our curriculum was designed to be taught entirely offline, relying on face-to-face discussions, printouts of photos and examples when needed, and zero technology. In many ways, therefore, the logistical aspects of our intervention were designed to mimic the prototypical learning environment of the average school-going child in rural India.

To ensure widespread acceptance of the program, Jeevika facilitated its promotion as an officially recognized government-certified course, thereby enhancing its credibility and increasing its external validity. Consequently, we were able to access and engage populations residing in remote rural areas that are typically underrepresented in studies on misinformation. Additionally,

[3]DataLeads received 400 applications from Bihar and selected around 50 teachers through an interview process, followed by a 2-day training to assess their skills. A key hiring criterion was a dynamic teaching style to foster discussion and engagement in the classroom. The final pool included school teachers, journalists, college professors, and fact-checkers, all based in Bihar. Teachers were assigned 6-9 classrooms across 2-3 districts, with each classroom having the same teacher throughout the intervention to maintain consistency.

[4]This is because students from multiple Gram Panchayats (GP) sometimes attend the same public school. GPs are groupings of villages governed by the same village council. As we note below in our sampling section, we determined that spillover potential between villages is high within the same GP but relatively lower across GPs.

the libraries that Jeevika provided access to were established facilities offering familiar environments for students and providing essential amenities such as seating arrangements, blackboards, and conducive classroom settings which allowed us a minimal level of standardization across classrooms, compared to public schools which have a larger variation in resources and problems with absenteeism.[5]

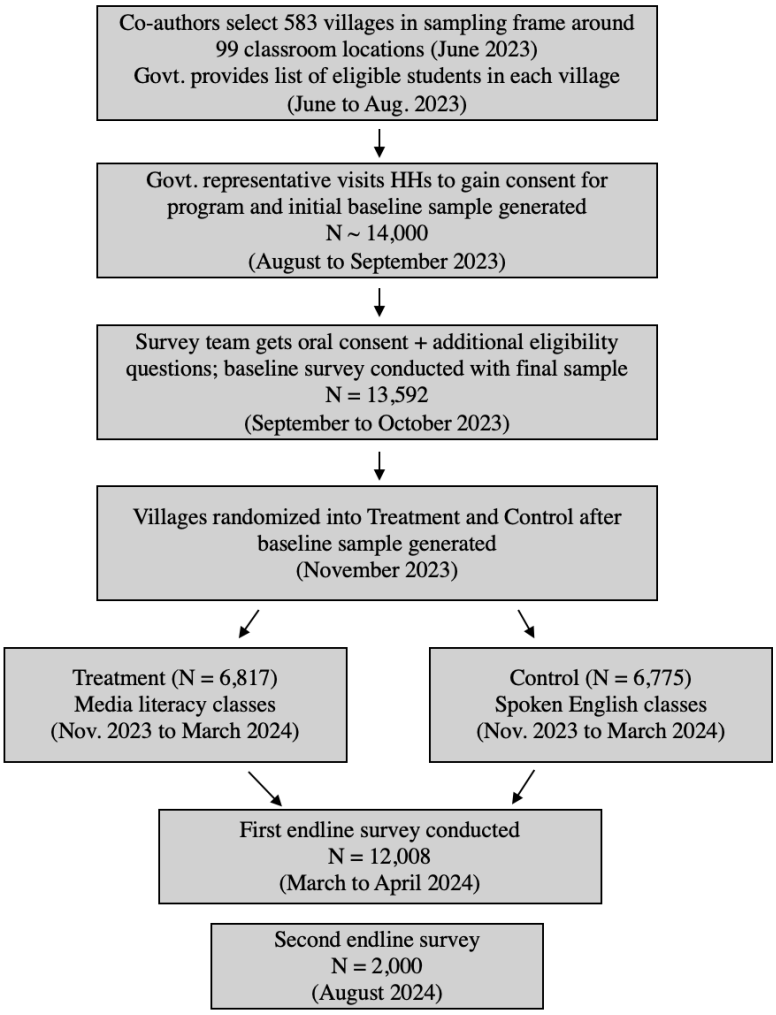
These efforts collectively aimed to incentivize enrollment and participation in the program. Notably, past studies using educational methods to combat misinformation, particularly in the Global South, have often reported null results (Badrinathan, 2021; Ali and Qazi, 2021; Blair et al., 2023). This aligns with a broader literature in comparative politics that suggests information provision alone may not sufficiently change beliefs and behaviors. Citizens often face constraints that hinder the translation of new information into action. Key among these is the need for citizens not only to receive novel information but also to be incentivized to prioritize the issue and believe in the efficacy of their actions (Kosec and Wantchekon, 2020). Furthermore, group dynamics can influence behavior; knowing that others are also informed may create social pressure that facilitates change (Lieberman, Posner, and Tsai, 2014). Therefore, in our study context, the educational classroom setting and the iterative nature of the intervention, featuring exposure to authoritative figures and peer interactions were designed to enhance receptivity to information interventions. Null effects in interventions may also result from insufficient support from governing authorities and disparities in implementation facilitation, particularly in developing countries where elite involvement is critical for mobilizing citizens, especially in rural areas (Rao, Ananthpur, and Malik, 2017). Partnered with Jeevika helped us with reducing these challenges. Additionally, enlisting external educators, conducting the intervention in well-resourced library settings, and implementing a comprehensive field monitoring plan to were all design choices aimed at increasing compliance and reducing attrition among educators and students (details on monitoring and compliance are provided in later sections).

[5]Data from the ASER survey, the Annual Status of Education Report which provides data from annual surveys on children's schooling and learning levels in rural India, highlights some of these issues. For example, their 2022 report points out that on the days that ASER surveyed schools, only 50% of enrolled children were actually present in public schools in Bihar; that number is the lowest in the 28 states in India that the survey spans.

4.3 Sampling, enrolment and baseline data

Figure 1 provides an outline of the timeline and flow of recruitment and roll out of the study. We sampled villages in close proximity to each library and randomized roughly 50% of these around each library to receive our treatment; the remaining served as control villages. Appendix A provides detail on our sampling procedures which ultimately allowed us to select 583 villages which served as the sampling frame for this study, with approximately half of these villages assigned to treatment.

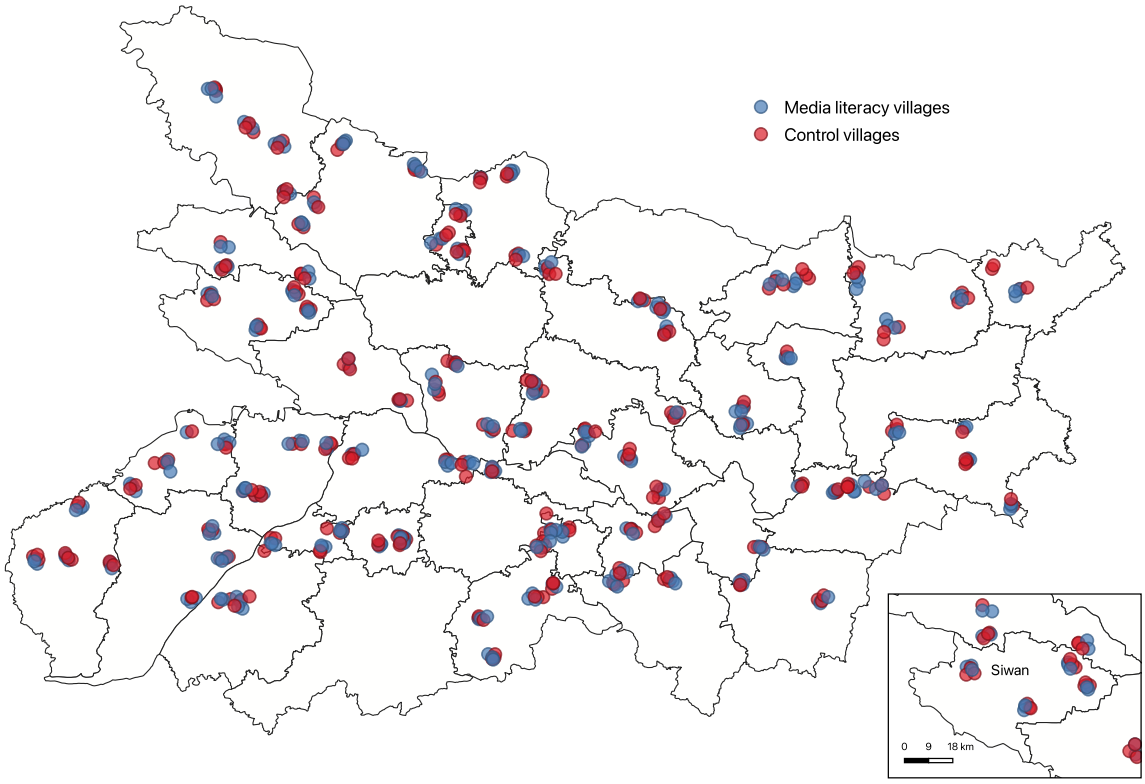
Figure 1: Flow, Timeline, Enrollment and Randomization



In each of the 583 selected villages, Jeevika provided a list of eligible students based on household data from enrolment in state programs. From this list, eligibility criteria for our study included households with children enrolled in government schools in grades 8 to 12. Jeevika staff

visited these households to confirm the presence of an eligible child, explained the program, and asked both the student and a parent or guardian if they were interested. From these visits, a final list of 20 to 24 interested students was created for each village. Figure Figure 2 shows location of treatment and control villages across Bihar.[6] Crucially, we note that randomization occurred *after* students opted in, avoiding issues with differential opt-in rates between treatment and control. Everyone involved in the study – including teachers, implementation partners, government officials, and coauthors – were blind to treatment status during recruitment and baseline data collection. During household visits, students and parents were informed only that they could participate in a free, government-offered certificate course with four sessions designed to benefit their future careers.

Figure 2: Map of treatment and control villages



Next, our survey team conducted the baseline survey in-person from September to Oc-

[6] Appendix A shows the locations of the 99 libraries where we held classes.

tober 2023. Enumerators visited homes on the Jeevika list, applying three additional eligibility criteria: obtaining oral consent from the student and a parent or guardian, assessing very basic (second-grade) reading comprehension in Hindi, and confirming that the student could attend all four sessions at their designated library location. Households failing any criterion were replaced in the sample. Once the baseline survey was complete, villages were randomized into treatment and control groups. Students remained unaware of whether they would receive media literacy classes or control until the first session. Appendix A provides more detail on sampling procedures.

The baseline survey collected demographic, household, and attitudinal data, including items on perceptions of the state, media usage, views on science and vaccines, and social ties. Our baseline sample included 13,592 respondents across 583 villages, with 49.9% assigned to treatment and 50.1% to control, 58% of whom were female.[7] In Appendix A, we show balance tables confirming that respondents in treatment and control groups were balanced on key demographics, attitudes, and behaviors, though we note a minor difference in average grade (age) level between groups. The Appendix also shows that treatment and control villages themselves were balanced in key demographic variables based on census parameters.

4.4 Control condition

Control group units participated in four modules of conversational English language classes, serving as a placebo rather than a pure control in order to achieve parity in effort exerted by students. We aimed to create comparable classroom dynamics and peer interactions, varying only the content of instruction. Encouraged by Jeevika, we additionally wanted control group respondents to gain in some way from the program and hence chose a topic that fostered engagement without confounding misinformation outcomes. Subjects like mathematics, science,

[7] Respondents ranged from grades 8 to 12, with the median being grade 10, and 96% were enrolled in government schools. The sample was 91% Hindu and 69% OBC, aligning with state demographics from census data. Language diversity included 43% from Hindi-speaking households, 30% Bhojpuri, and 9% Magahi. Most students came from households where fathers' median education was grade 6-9 and mothers' median education was grade 1-5. Household assets indicated socio-economic status of the sample: 15% owned a refrigerator, 3.6% owned a washing machine, and 19% had an active mobile phone with internet access. Attitudinally, 90% trusted newspapers, 84% trusted TV, and 61% trusted social media. Additionally, while 77% had been vaccinated for COVID-19, 87% believed in alternative medicine like ayurveda and homeopathy.

history, and geography were excluded due to their redundancy with the standard curriculum or their association with national identity narratives. Alternative hands-on topics (like cooking or knitting) were also discarded due to gendered uptake concerns.

We ultimately implemented a curriculum of four sessions focused on basic conversational English to enhance students' future career prospects. The modules targeted students with limited English exposure and emphasized spoken skills over reading and writing. The curriculum highlighted the importance of English proficiency for career opportunities, covering self-introduction, naming common objects, describing activities, and formulating questions. Activities mirrored those in the treatment group, including role-playing and group exercises, while avoiding media, technology, or politics. Due to the basic nature of the instruction, students likely would not gain sufficient proficiency to navigate new information sources independently.

The teacher selection and training differed between the treatment and control groups. Treatment group teachers were recruited and trained by DataLeads, while English class teachers were sourced from a local Bihar consultant, resulting in variations in socio-economic characteristics and teaching experiences. Consequently, the treatment effects we measure are influenced by both the treatment content and the teachers' differing backgrounds. Demographic characteristics of the teachers by treatment condition are summarized in Appendix F.

4.5 Endline data and compliance

Our first endline survey was conducted in the weeks following the end of the fourth and last session. This took place face-to-face in respondents' homes, by the same enumeration team that carried out the baseline. Because of the large sample size, the endline was conducted 1 to 5 weeks after the last class of the module, depending on the location of the village. Although we did not randomize the order in which villages were surveyed at endline, we expect that endline survey timing, while dependent on a number of village-level covariates, is independent of treatment status. At the first endline, we were able to re-contact 12,008 of the total houses sampled at baseline, with an attrition rate of 11.3%. There is no significant difference in attrition between treatment and control groups, although we do find that attrition is lower among female students and those with higher parental levels of education, and higher among those in higher school

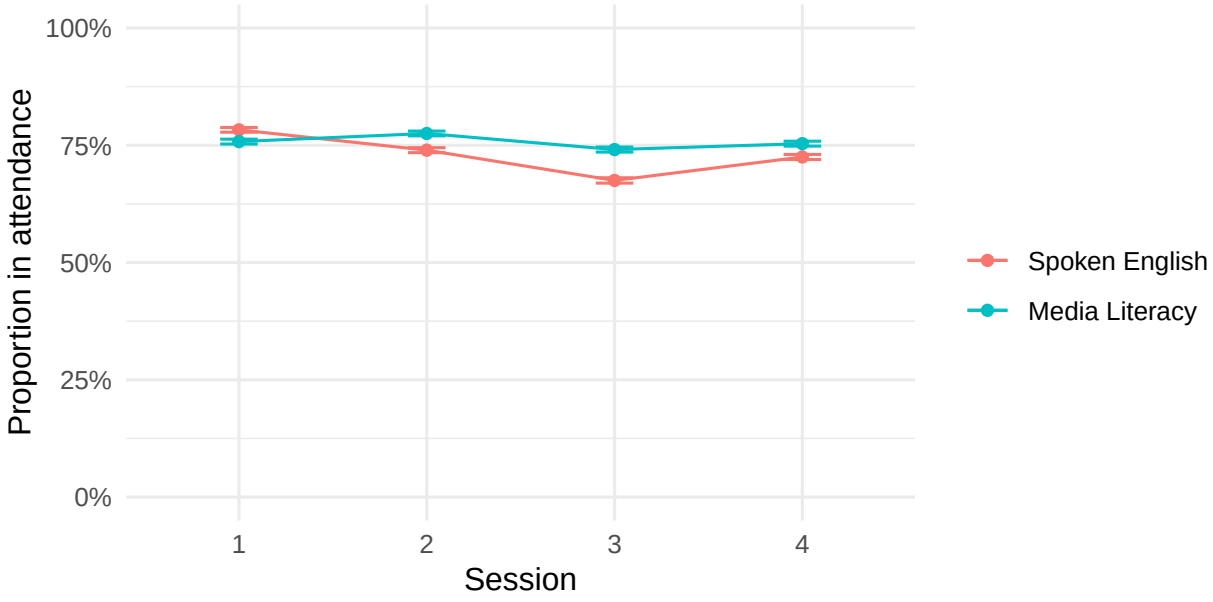
grades (see Appendix D). Moreover, from fieldwork and interviews with enumerators, we note that houses that attrited at endline did so because we were unable to contact them after several tries (in most cases, this was because the respondent or parent was not at home). Crucially, no household refused our survey team entry for the endline survey. We conducted a second endline survey about 4 months, on average, after the intervention, to assess if treatment effects persisted. This survey was conducted over the phone with a random subset of 2,000 students and one parent/guardian.[8]

To ensure compliance, we implemented a detailed monitoring system. For students, Jeevika staff, known as *didis*, regularly reminded households about upcoming classes, and they were compensated for their efforts. Students were motivated by the promise of a government-issued certificate upon completing the program. They also received notebooks, pens, flyers, homework, and snacks during each session. For teachers, external monitors made random visits to verify their presence, documenting observations through brief surveys. Co-authors also visited during initial and final sessions. Teachers were required to upload attendance data and classroom photos via an app, ensuring adherence to the schedule.

On average, across both treatment and control students attended 2.97 classes and 52.7% of the sample attended all four classes. This attendance allows us to measure whether students were fully non-compliant (attended 0 sessions), fully compliant (attended all 4 sessions) or partially compliant (1-3 sessions). Crucially, our definition of attrition does not hinge on compliance with the treatment protocol. Even participants who are (partially) non-compliant with the treatment but continue to engage with the study by completing the endline survey are not classified as having attrited. Our compliance data demonstrates that on average, 76% of the sample attended the first class. Crucially, we detect no significant difference in attendance across treatment and control, with similar proportions attending both sets of classes. However, we do see a significant drop off in attendance for control group respondents during session three, through we note that the difference is not substantively very large (67% in control group and 74% in treatment) and dissipates during session 4. Figure 3 graphs attendance data by treatment group and by session.

[8]The time gap between the first and second endline surveys varied across households because it took about 30 days to survey all homes in each round. For some respondents, the gap was around 3 months, while for others, it extended to 5-6 months. Therefore, we report an average gap of 4 months.

Figure 3: Compliance data across treatment and control



Further, we find women (0.37 more sessions attended) and younger students (0.13 fewer sessions attended per additional school grade) were significantly more likely to attend sessions, compared to men and older students (see Appendix D). We do not detect any differences in compliance with attending sessions as a function of internet use, household assets (income), or trust in the media.

4.6 Outcome measures

We pre-specified seven primary families of outcomes for the first endline survey: awareness of misinformation, accuracy discernment, sharing discernment, health attitudes, trust in sources, demand for fact-checking, and engagement with misinformation efforts.[9] For each family of outcomes, we compute inverse-covariance weighted (ICW) indices that combine and weight individual items, that are standardized relative to the control group mean and SD.[10] The specific variables and procedure for constructing the index were pre-specified (see Appendix E). We describe our primary outcomes as we present our main results, in the next section. In the second

[9]Our pre-analysis plan (PAP) was posted to OSF at the beginning of endline data collection in February 2024. It specifies primary and secondary outcomes and how they are constructed, main analyses, heterogeneity analyses, and robustness tests. The PAP is available here: https://osf.io/h43qn.

[10]For details on the rationale behind using ICW indices and how exactly we construct the outcome measure, see Appendix E.

endline survey, we measure two of the seven outcomes – accuracy discernment and health attitudes – for both respondents and one parent/guardian, alongside self-reported items seeking to elucidate mechanisms.

5 Estimation and Results

Due to the possibility of non-compliance, our main specification estimates the intent-to-treat ITT_Y effect: the effect of being assigned to the treatment group. To test hypotheses about the direct overall effect of the treatment on average outcomes, we use the following two models:

$$Y_{ijk} = \beta_0 + \beta_1 T_i j + \sum_{k=1}^{m-1} \gamma_k + \varepsilon_i \tag{5.1}$$

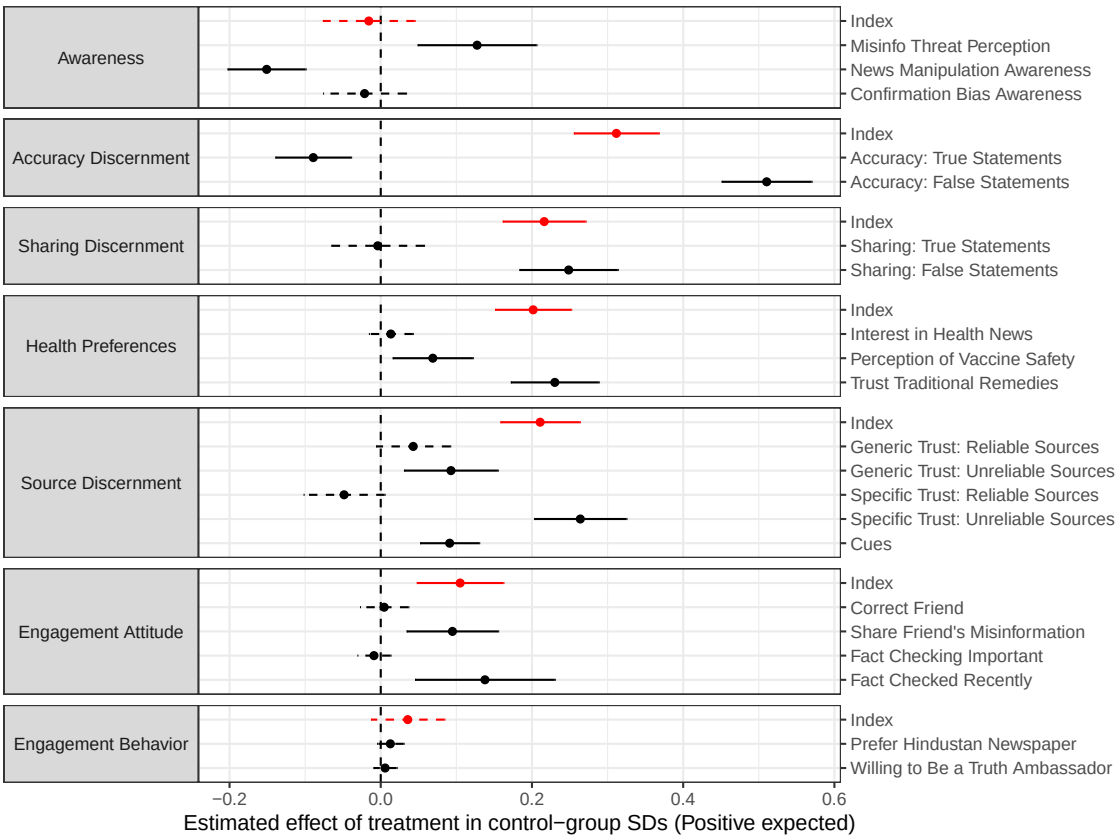
$$Y_{ijk} = \beta_0 + \beta_1 T_i j + \sum_{c} \alpha_c X_{ci} + \sum_{k=1}^{m-1} \gamma_k + \varepsilon_i \tag{5.2}$$

where Y_{ijk} is the primary outcome of interest Y for student i in classroom j and district $k \in \{1, ..., m\}$, β_0 is the intercept, $T_i j$ is a treatment indicator, α_c denotes the coefficient for the control variable X_c, γ_k denotes fixed effects for each district k, and ε_i denotes the random error term for individual i. β_1 denotes the estimated effect of treatment assignment (ITT) on outcome Y. To estimate this equation, we use linear regression with heteroskedasticity-robust standard errors, clustered at the classroom or village level.

5.1 First Endline

We first examine the overall effect of BIMLI by comparing endline outcomes from the first endline survey for those in the treatment classes with those in the control group. Because we calculate ITT_Y, the estimates of treatment effect we present here can be seen as conservative because of dilution due to partial non-compliance. Our main results are summarized in Figure 4 which shows the estimated effect of assignment to treatment on our seven pre-registered families of outcomes. We find that the treatment positively affected five of the seven outcomes and detail these results below. Tabular results are available in Appendix H.

Figure 4: Estimated effect of assignment to BIMLI treatment



Notes: This figure plots the effect of BIMLI for 7 outcome families. Each index is an ICW calculation of components within an outcome family. Each component is standardized relative to control mean and SD. Confidence intervals are based on standard errors clustered at the village (classroom) level and tabular results are in Appendix H.

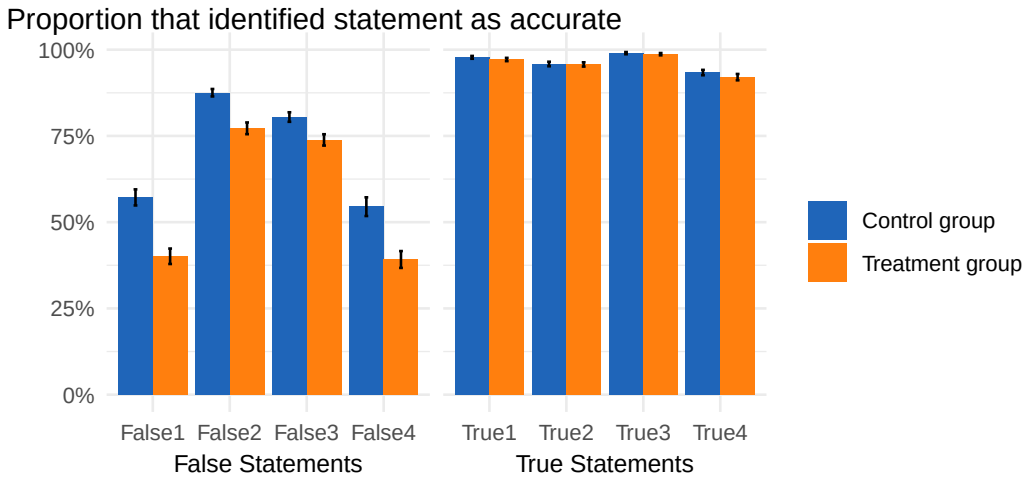
5.1.1 Accuracy and sharing discernment

Recent years have seen a convergence in academic work around making one measure central to testing whether misinformation interventions are effective: discernment between true and false information. Focusing on the premise that misinformation counter-measures should help citizens maximize the accuracy of their beliefs and of the content that they share, this measure requires (1) a design in which respondents rate a mix of both true and false content, and (2) an analysis that includes examining discernment between the two (Guay et al., 2023). Consequently we adopt a standard discernment measure consistent with the misinformation literature by asking respondents to rate their perceived accuracy, on a 4-point scale ranging from very accurate to not at all accurate, of a set of 8 news stories balanced on veracity. Crucially, we note that only 1 of 8 of these news stories was an example discussed in the classroom: teachers talked to students

about Covid-19 vaccine safety as that topic was a part of the curriculum but did not discuss any of the other issues. Thus the discernment effects we estimate are largely a function of skill and application, rather than mere recall. We then used the same set of 8 stories to measure sharing discernment, asking respondents to state their intention to share a given headline, consistent with literature (Pennycook et al., 2021). Since some previous work has shown that thinking about the accuracy of a story can affect intentions to share, we randomize the order of the sharing and discernment battery such that one half of the sample is asked each set of questions first.

The selection process for dependent variable stories was driven by key factors. We prioritized stories respondents were likely to be already familiar with to avoid introducing new misinformation and to ensure cultural relevance, maximizing the potential impact on beliefs. After extensive fieldwork and piloting, we identified the most commonly believed health-related misinformation. Each chosen story had been independently debunked by at least one fact-checking service in India. Headlines were presented to respondents in random order.

Figure 5: Accuracy discernment by headline



With respect to accuracy discernment, our results show that our treatment was overall highly successful in significantly helping respondents discern between true and false stories (Figure 4). The magnitude of this effect, a 0.32 SD increase in discernment relative to the control group, is substantively large relative to other comparable work. For example, Guess et al. (2020) find that their digital literacy intervention in India led to a 0.11 SD increase in discernment, while Gottlieb, Adida, and Moussa (2022)'s intervention in Cote d'Ivoire produced effect sizes 0.12 to

0.15 SD. Further, we see variation in the true and false components of the discernment measure: we find that while the treatment made respondents marginally more sceptical of all news, the magnitude of this effect is much larger for the false statements, and hence the overall effect is a net positive. In Figure 5 we break down the true and false components into individual headlines. The graph demonstrates that large proportions of respondents in the control group thought that false statements were accurate and the treatment significantly decreased respondent's perceived accuracy of all 4 false stories, with effect sizes ranging from 0.44 SD (cow urine can cure Covid) to 0.19 SD (mobile phone towers cause cancer). With respect to true stories, there is little variation in how treatment and control group respondents rated these stories; on average all respondents were better at discerning true stories relative to false.

With respect to sharing discernment, we find that the treatment has an overall significant effect on affecting respondents' willingness to share each story we asked them about (0.22 SD). Overall our results on discernment confirm that the treatment was successful at helping respondents prioritize accuracy when believing content as well as sharing it. That we are able to detect effects on stories that were not discussed in the classroom demonstrates a crucial learning component that treated respondents were able to glean from the program. Further, unlike previous studies on misinformation that measure outcomes immediately after treatment, or even as part of the same instrument, given the gap between classroom sessions and the endline survey we can be confident that recall or demand effects are not primarily driving this finding.

5.1.2 Health preferences

At endline, we measured health preferences through three components: interest in health news, vaccine safety perceptions, and reliance on alternative medicine. Respondents rated their interest in health news on a scale from very interested to not interested. For vaccine safety, they rated the safety of both the COVID-19 and chickenpox vaccines. To assess reliance on alternative medicine, respondents were asked if they would visit a traditional healer or use home remedies for serious illnesses, and whether they agreed that ayurveda and homeopathy could cure serious diseases. At baseline, we observed significant variation in responses to these questions. Over 77% reported being vaccinated against COVID-19, with even more expressing intent to get vaccinated, likely

due to government-led campaigns. However, 87% also believed in the effectiveness of alternative medicine like ayurveda and homeopathy. Many respondents in fieldwork discussions mentioned relying on health advice from friends and family rather than doctors, with home remedies and medical misinformation being common, even for serious illnesses.

Despite the prevalence of health misinformation and reliance on alternative medicine, we show that BIMLI was able to significantly alter respondents' health preferences overall (index 0.21 SD). Item-wise results indicate that while the treatment reduced vaccine hesitancy and significantly reduced stated reliance on alternative forms of medicine to address serious health issues. This finding holds significance for several reasons. First, traditional home remedies and the misinformation surrounding them have long existed in the country, passed down through generations, suggesting that these beliefs may be deeply ingrained and therefore resistant to change. Additionally, prior research has indicated that belief in medical misinformation in India is associated with other identities such as religion and partisanship, and given that these identities underpin enduring societal divisions (Chauchard and Badrinathan, 2024), motivated reasoning may impede the effectiveness of misinformation countermeasures (Taber and Lodge, 2006). Despite this backdrop, our findings demonstrate that BIMLI has a significant impact on altering respondents' health preferences, including their preferences regarding which forms of medicine to rely on during crises, beliefs about the efficacy of alternative treatments, and their ability to discern between true and false medical misinformation.[11]

5.1.3 Trust in sources and source discernment

We introduced a series of novel items to assess how respondents navigate trusting and discerning between various news sources. Typically, trust in mainstream news is measured by classifying outlets like CNN or Fox News as trustworthy or not. However, we avoided this approach for two reasons. First, due to recent political developments, distinguishing between trustworthy and untrustworthy media in India is difficult, as most mainstream outlets have engaged in polarizing

[11] Of course, changing these preferences does not guarantee that better, more scientific forms of medicine and/or medical treatments are necessarily accessible for and available to our respondents or to their family. We nonetheless believe that shifting such preferences is an important outcome in the short run, insofar as it should allow respondents to become more discerning in their consumption of health-related misinformation, and an important political outcome in the long run, in the sense that it may help shape citizens' demand for better services.

content and misinformation (Amanullah, Nadaf, and Neyazi, 2023; Mohan, 2021). Second, respondents heavily rely on informal sources, such as word of mouth and local elites (Gadjanova, Lynch, and Saibu, 2022). Thus, our source discernment index includes three components.

First, we measure general source discernment by asking respondents to rate their trust in sources, including transmitters (e.g., word of mouth), mediums (e.g., radio, Facebook), and institutions (e.g., WHO). For our index, we focus on four sources: two we expect to increase trust in (MBBS doctors, ASHA healthcare workers) and two we hope to decrease trust in (ayurvedic doctors, rumors). Next, we assess situation-specific trust by using a vignette where respondents seek emergency advice for a sick family member, rating their trust in various sources. We provide three trustworthy sources (community health center, government materials, TV doctors) and three untrustworthy ones (family remedies, WhatsApp forwards, TV interviews with ayurvedic doctors). This helps distinguish between general and specific trust and separates transmitters from mediums. Finally, we explore which factors foster trust in information encountered on social media, examining signals like likes/shares and the sender's community, as well as message tone and emotionality. Our results show that BIMLI significantly changed how respondents interact with and trust sources, with a notable shift in the index (SD = 0.21). The treatment primarily reduced trust in unreliable sources rather than increasing trust in reliable ones and also reduced reliance on untrustworthy cues online.

5.1.4 Engagement with misinformation

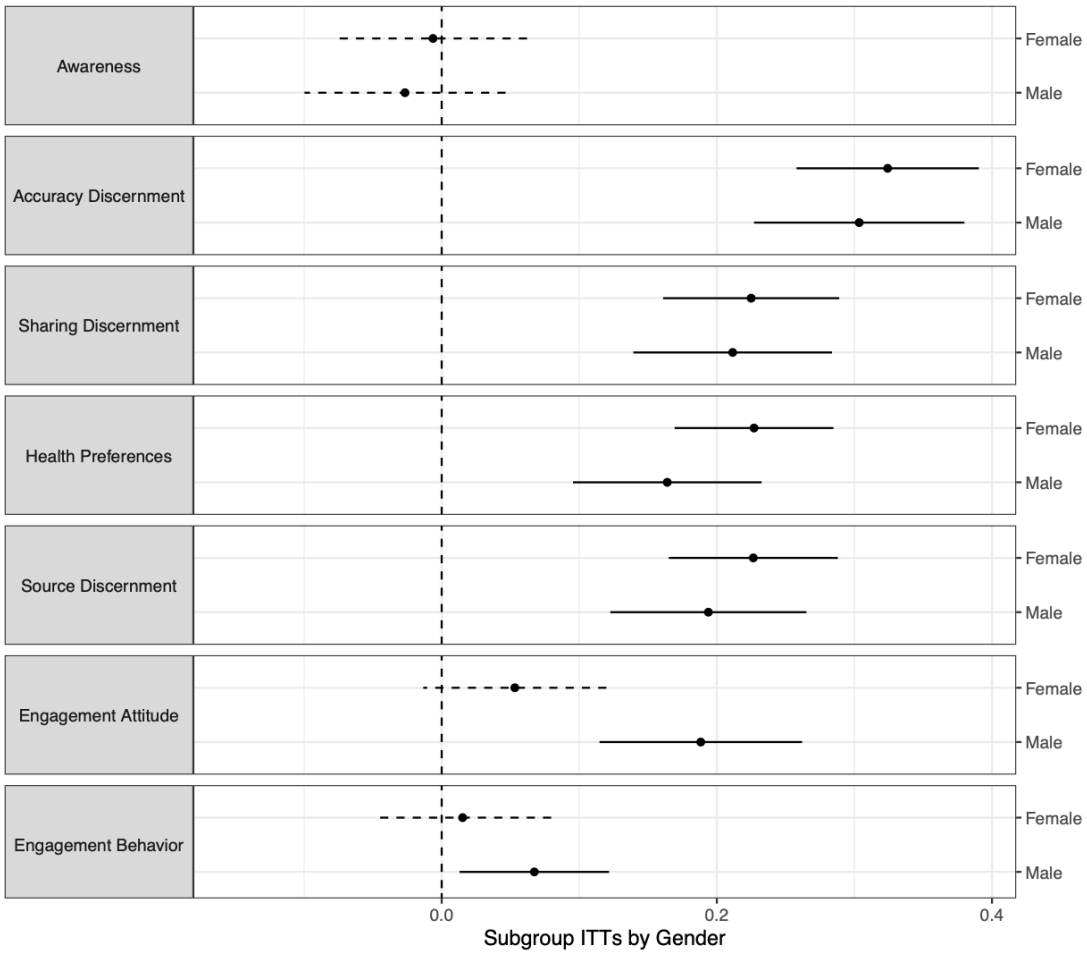
We measured engagement with misinformation through both attitudinal and behavioral measures. On attitudes, we aimed to shift norms around perceptions of misinformation by asking about demand for fact-checking with four self-reported measures. These ask: (1) whether respondents would correct a friend they know is sharing misinformation and (2) themselves share misinformation from friends when they know stories are incorrect, (3) We measure self-reported importance of verifying and fact-checking information before sharing on a 3-point importance scale. (4) We measure the self-reported frequency with which respondents say they have tried to verify information in the last 2 months. The results show the treatment significantly influenced respondents' attitudes toward misinformation, though item-level variation exists. Treated

respondents reported a greater likelihood of abstaining from sharing misinformation, even from close acquaintances, but were reluctant to correct it. This aligns with research suggesting that cultural factors in India, like respect for social norms and relationships, discourage direct confrontation over misinformation (Malhotra and Pearce, 2022). While respondents hesitated to correct friends, the shift toward not sharing misinformation suggests that the treatment was effective. Regarding fact-checking, a ceiling effect was observed, with 83% of respondents, including 82% in the control group, already affirming its importance, limiting the treatment's effect on this belief.

We included two behavioral measures. First, we asked respondents to enter a lottery to choose between two free subscriptions: a year-long subscription to *Hindustan*, a reliable Hindi newspaper, or *Manohar Kahaniyan*, a popular magazine featuring stories of horror, mystery, and romance. We expected higher demand for news in the treatment group. Second, we invited respondents to become truth ambassadors, tasked with helping dispel misinformation during future crises. We anticipated higher willingness to engage in this role among the treatment group. We show results in Figure 6. The results indicate that, on average, the treatment did not significantly impact these behaviors, as the overall index effect is null, and item-level responses show no deviation from zero. However, this average null effect obscures important gender-based variation. When examining intent-to-treat effects by gender, we find significant heterogeneous treatment effects on misinformation engagement measures, despite no other dependent variable indices exhibiting gender variation. Specifically, boys are more likely to report an intention to engage in misinformation countermeasures, both in attitudes and behaviors, while the treatment did not affect women on these items.

This result is consistent with India's patriarchal context, where strong gender norms favor men (Brulé, 2020; Carpena and Jensenius, 2021; Prillaman, 2023). The measures, particularly those related to behaviors and behavioral intentions, reflect not only attitudes towards misinformation but also the capacity and willingness to perform community-based actions, which may require shifts in gender norms, safety for women, and permission to engage publicly. For example, correcting a friend's misinformation may require assertiveness and confrontation—traits not directly targeted by the intervention, and particularly difficult to move with women. Thus, while both girls and boys showed equal improvements in trust in reliable news sources and the

Figure 6: Effect of assignment to BIMLI by Gender Subgroup



Notes: This figure plots the effect of BIMLI for 7 outcome families, each with subgroup coefficients by gender. Each index is an ICW calculation of components within an outcome family. Each component is standardized relative to control mean and SD. Confidence intervals are based on standard errors clustered at the village (classroom) level.

ability to discern truth from falsehood, behavioral change appears more challenging when deep-rooted cultural and gender norms serve as barriers. Boys, in contrast, are more likely to report a willingness to change their behavior.

5.1.5 Awareness

To assess awareness of misinformation and recall of classroom material, we developed an ICW index with five items. The first measured perceptions of misinformation as a threat, rated on a 1 to 5 scale against issues like unemployment or corruption. Although exposure to BIMLI significantly heightened perceptions of misinformation as a threat, 78% of respondents already

viewed it as such, limiting further shifts. We also measured awareness of media and cognitive biases using items adapted to the Indian context from Ashley, Maksl, and Craft (2013). On cognitive bias, measured using items such as "People pay more attention to information that fits their beliefs," treated respondents showed no improvement over the control group (p = 0.7), possibly due to the time gap between lessons and the survey. However, significant effects on discernment suggest that respondents did apply classroom lessons, even if recall was difficult. Treated respondents were less likely to recognize media bias, unexpectedly disagreeing with statements like "News organizations select stories to attract larger audiences." This may reflect confusion over what news organizations should do versus what they actually observe.

5.1.6 Heterogeneous treatment effects

Finally we discuss heterogeneous treatment effects. We pre-registered analyses based on pre-test variables identified in the literature as significant in this context. To proxy motivated reasoning, we assess interaction effects with partisan identity and attitudes towards non-scientific belief systems (e.g., Ayurveda). We also examined household mobile internet access to see if connectivity influenced responses, and pre-test knowledge of basic science. emographically, we consider age (class grade of students) and gender. Additionally, to ensure that the treatment did not differentially affect households in high and low spillover strata, we estimate interactions by spillover stratum. The results, detailed in Appendix J, show no consistent patterns. Notably, departing from the gender subgroup effects already reported above, we found no interaction effects for mobile internet usage, age, or partisanship, despite past research showing that partisanship can dilute the effectiveness of misinformation interventions (Flynn, Nyhan, and Reifler, 2017; Badrinathan, 2021). These results suggest that belief change in our context was not influenced by prior identities or internet access, and that our treatment was robust enough to counteract motivated reasoning.

We observe some interaction effects with science knowledge, where the treatment appears to help those with higher science knowledge improve source discernment and engagement attitudes. Prior attitudes toward Ayurveda also interact with the treatment on engagement attitudes. However, these effects are small, inconsistent, and have large confidence intervals, suggesting

they may be due to chance. Lastly, we find no differences in treatment response between high and low spillover strata.

5.2 Second Endline

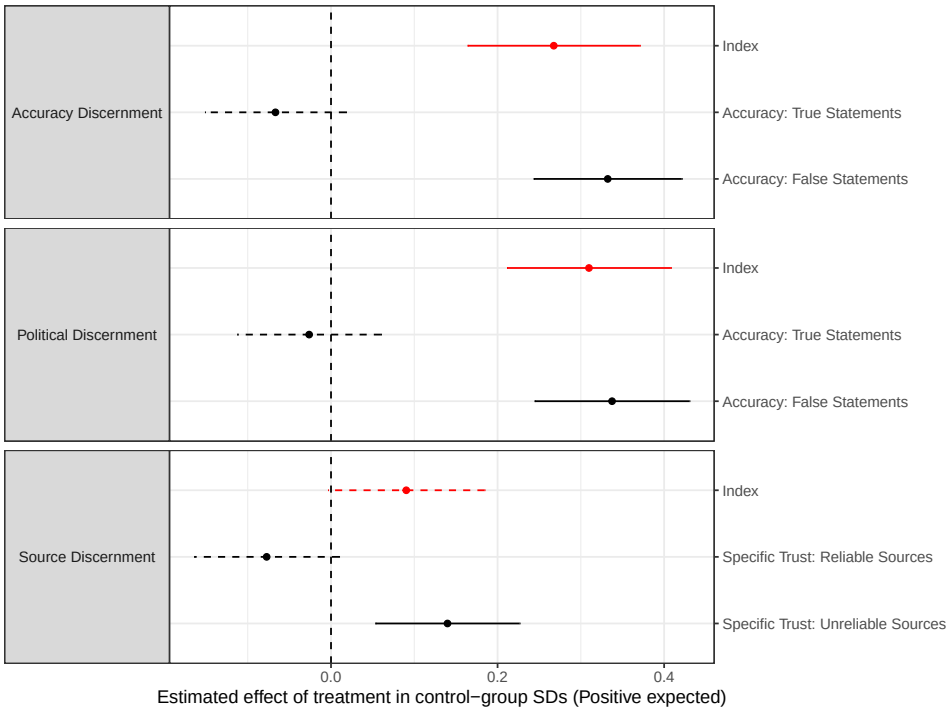
We conducted a follow-up survey with a random subset of 2,000 respondents approximately four months after the intervention to assess its long-term effects and potential network impacts. The extended time gap is particularly relevant, as India's 2024 general elections occurred between our two endlines—a period when political and partisan attitudes typically become more entrenched (Michelitch and Utych, 2018). The follow-up had three main objectives: (1) to evaluate whether the ability to discern true from false information persisted over time, (2) to assess if respondents could apply this skill to political stories, a new set of stories and a domain not covered by the intervention, which focused solely on non-political topics, and (3) to examine potential network effects. Since students were encouraged to discuss what they learned with their families, we also interviewed one parent or guardian per respondent to determine whether families of treated students demonstrated improved discernment abilities.

Remarkably, our findings indicate that participants in the treatment group continued to exhibit an enhanced ability to distinguish between accurate and inaccurate health information (0.27 SD), and effect sizes remained substantial (Figure 7). Additionally, we find that treated respondents demonstrated a significantly greater capacity to accurately identify the veracity of political stories (0.31 SD). This outcome is especially noteworthy given that our intervention was exclusively focused on health-related content and did not address political topics. Moreover, the political stories included in this survey were entirely new and had not been assessed in the initial endline. Nevertheless, treated respondents were able to effectively discern truth from falsehood on political topics. This finding carries significant implications: education about misinformation, even when concentrated on a specific domain (in this case, health), has the potential to transfer its benefits to other domains, and its effects can endure over a substantial period of time.

From our interviews with parents of respondents as part of this second endline, we find that parents of treated students were significantly better at discerning true from false health information (0.29 SD), as demonstrated in Figure 8. This result is particularly notable as it highlights

the potential for "trickle-up" socialization, where children's learning influences their parents' attitudes and behaviors (Dahlgaard, 2018). It also suggests that sustained learning may generate valuable network spillover effects. While the exact mechanisms of parental learning—whether through handouts, conversations with their children, or other means—remain unclear, the significant change observed in parents' ability to discern true from false information is noteworthy.

Figure 7: Effect of assignment to BIMLI treatment on 4-month follow up



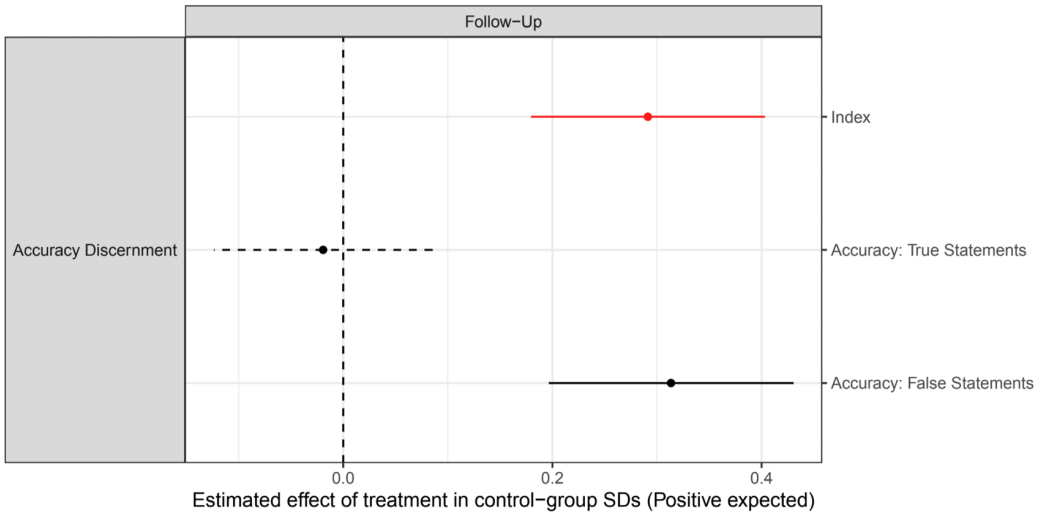
Notes: This figure plots the effect of BIMLI amongst a sub-sample of respondents interviewed 6 months post the intervention.

5.3 Mechanisms

We now discuss the mechanisms underlying our findings. As hypothesized in section 2, treatment group respondents may have shown changes in attitudes and behavior due to at least three distinct mechanisms: they could have reiterated classroom learnings (*recall* mechanism) or applied new strategies and skills acquired in class (*application* mechanism); alternatively, they might have been influenced to perceive misinformation or to express themselves with regards to this issue differently (*norms* mechanisms).

To evaluate these various possibilities, we asked treatment group respondents during the

Figure 8: Effect of assignment to BIMLI on treatment group parents



Notes: This figure plots the effect of assignment to BIMLI on parents of treated respondents (versus control).

second endline survey to describe how they believed the treatment affected them. They selected from three options: whether it provided new information, corrected existing misconceptions, or socialized them into adopting certain beliefs (such as refraining from sharing misinformation in public). We found that a significant proportion—40 percent—indicated that the treatment imparted new information. Meanwhile, 29 percent felt it corrected existing beliefs, and 32 percent believed it changed norms, suggesting that all three mechanisms likely contributed.

To investigate further, we presented respondents with a scenario involving someone stating misinformation and asked how they would primarily respond. Treatment group respondents were significantly more likely to indicate they would demonstrate a strategy to verify information in the future, as well as reinforce not sharing misinformation. This suggests that the mechanisms of learning and applying skills, as well as shifting norms, were both at play.

Lastly, we emphasize that the follow-up survey focused on political stories—topics not discussed in the classroom or during the treatment. The significant effects observed in this area imply that the mechanism at play goes beyond mere recall or changes in social norms about what respondents believe is acceptable to express publicly. Discernment requires specific skills and tools, particularly in unfamiliar domains. This supports the notion that the primary mechanism through which BIMLI operated was the provision of new tools and skills.

6 Discussion and Conclusion

In this study, we evaluated the impact of a large-scale, grassroots classroom-based intervention aimed at combating misinformation, implemented among over 12,000 adolescents in Bihar, India. In collaboration with a local government body, we developed a media literacy program that spanned half a school year. Intent-to-treat estimates revealed a strong, significant improvement in students' ability to process and critically assess information, as well as apply classroom lessons to real-world situations. By the end of the program, treatment respondents exhibited greater discernment in evaluating information and deciding whether to share news items. The intervention also shifted health preferences, reducing reliance on alternative medicine, and improved students' ability to assess the credibility of sources. These effects persisted among a subsample interviewed 4 months later. Importantly, follow-up assessments showed that students were able to accurately discern political news, a topic not covered in the program, demonstrating the transferability of the skills learned. Finally, we found that parents of treated students were significantly better at distinguishing true from false information, indicating that such educational interventions can have spillover effects within social networks, with knowledge trickling upwards through socialization.

These findings are particularly important given the challenges many media literacy interventions face, often yielding mixed or null results (Blair et al., 2023). Our intervention showed significant effectiveness among a difficult sample in a challenging context. Bihar, where the study was conducted, has low prioritization of education and a high dropout rate (Muralidharan and Prakash, 2017), with 42% of students leaving school before 10th grade. Although we achieved around 70% session compliance, the lower-than-ideal rate may reflect the realities of working in rural areas with limited state capacity. Bihar consistently underperforms in various national metrics due to extreme state weakness, so it was not readily foreseeable that a curriculum like BIMLI would have such notable effects. Despite the difficult setting and politicization of health issues in India, adolescents engaged with the program, which influenced their preferences, behaviors, and information consumption. While adults might have been more resistant to change due to motivated reasoning (Taber and Lodge, 2006; Badrinathan, 2021), this supports our argument that interventions should target populations more open to change—like young people seeking

education in underserved areas.

Despite these encouraging findings, we recognize several limitations of the study. First, the intervention was delivered as a bundled treatment covering multiple components, making it difficult to isolate which aspects drove the results. For example, we cannot easily determine whether the focus on cognitive biases or the interaction with adults had a greater impact. Our intention was to design a comprehensive, high-dosage intervention, especially since many other media literacy programs have shown limited success. However, future iterations or extensions of this work could unbundle the curriculum to identify where the effects are most concentrated. Another limitation is the high cost and effort required to set up such an intensive and sustained intervention, raising the question of whether this dosage of treatment is necessary to combat misinformation. Indeed, many studies use brief fact-checks or primes and nudges and find effects. To explore this, we analyzed session attendance and interacted compliance with the treatment to assess whether more engagement altered outcomes. Results in Appendix K show that attending just one session had no significant effect on any outcomes, aligning with research suggesting one-time media literacy treatments are insufficient in polarized settings (Badrinathan, 2021). However, we found that attending two sessions produced significant effects on accuracy and sharing discernment, but no further gains with three or four sessions. In contrast, health preferences, source discernment, and attitudes toward counter-misinformation measures only shifted after three sessions.

These findings suggest two key points: First, discernment tasks are more easily influenced but require at least two sessions (or three contact hours) for significant change, with effect sizes two to three times larger than those in existing research. Second, outcomes requiring three or more sessions typically involve deeply ingrained attitudes, such as long-held beliefs in Ayurveda, which are harder to change than newer misinformation narratives. Research on polarization and social contact further supports that attitudinal changes are difficult despite intensive interventions (Scacco and Warren, 2018; Mousa, 2020). Our findings show that while challenging, even deep-rooted attitudes can change through sustained, education-based efforts. This highlights the need to move beyond light-touch interventions and recognize the importance of sustained engagement for meaningful attitude shifts, particularly in developing contexts.

Finally, the generalizability of our findings depends on the scope conditions we identify.

As stated earlier, our study sample is unique, characterized by low state capacity, limited access to credible news, and low socioeconomic status. To make the intervention work in this challenging context, we made specific design choices, such as bringing in external teachers to deliver the curriculum and partnering with a well-known state agency for rollout. For example, over 95% of parents surveyed, both in the treatment and control groups, said they would send their children to such a program again. When asked why, most (65%) said they wanted their children to learn, but a notable portion (30%) mentioned that they trusted Jeevika, the state agency. We thus acknowledge that one reason the program may have been successful is that it stood out in this context—whether due to its quality, teaching style, or being a rare high-quality educational opportunity. As a result, we cannot be sure if: (1) the intervention could work or even be implemented within existing public systems in India or Bihar—we concede this uncertainty; or (2) whether it would be effective for children in these other systems, who may have different characteristics than our sample. While we recognize this limitation, we also note that our evidence shows minimal variation in effects based on income, socioeconomic status, or political variables such as party identification, suggesting that the intervention could have similar impacts across different populations (see Appendix J).

That said, many contexts similar to ours exist across the Global South, where misinformation is prevalent due to low state capacity, limited access to credible news, and socioeconomic inequalities. In these regions, misinformation often spreads through informal networks rather than online platforms (Gadjanova, Lynch, and Saibu, 2022). The challenges we observed in Bihar—such as reliance on unverified sources and weak educational systems—are widespread. Thus, our findings may have broader relevance, suggesting that grassroots, education-based interventions like ours could be effective in other regions facing similar conditions, where misinformation thrives and reliable information is scarce.

Despite these limitations, our positive findings offer hope and carry significant implications for both academic research on misinformation and policy development. First, researchers should focus on sustained, iterative treatments, even if they are intensive and costly. In many settings, these may be the only viable solutions, especially where populations lack internet access, rendering platform-based solutions like fact-checking and algorithmic changes ineffective. From a policy perspective, modules like ours should be integrated into school curricula. When edu-

cational quality is high, such interventions can be seamlessly incorporated into existing school structures, as seen in Finland and New Jersey. Additionally, while we observed positive effects on parents, there is potential for even greater impact through siblings, friends, and others in respondent networks. This highlights education's capacity for network spillover and equilibrium effects that are often overlooked in other approaches.

We attribute these hopeful findings to the sample and setting we study: classrooms and schools have consistently been identified as pivotal sites for knowledge acquisition beyond the household, and public education systems play a crucial role as agents of socialization. Therefore, our study not only contributes to literature on persuasion and information processing but also examines the enduring impacts of education and learning. This aligns with existing work exploring the transformative potential of education within schools. Previous studies have investigated the use of education to reshape gender attitudes in India (Dhar, Jain, and Jayachandran, 2022) and foster nation-building efforts (Bandiera et al., 2019), along with the potential of interaction with the state via education to shape economic views (Davies, 2023). Further, scholars have explored the efficacy of educational tools such as textbooks in persuasion and attitude change (Cantoni et al., 2017), as well as their role in shaping perceptions of representation and marginalization (Haas and Lindstam, 2023). By situating our study within the broader context of educational interventions, we contribute to scholarly understanding of the multifaceted impacts of schooling on attitudes and behaviors.

References

Ali, A., and I. Qazi. 2021. "Countering misinformation on social media through educational interventions: evidence from a randomized experiment in pakistan.".

Amanullah, Arshad, Arif Hussain Nadaf, and Taberez Ahmed Neyazi. 2023. "Constructing the Muslim "other": A critical discourse analysis of Indian news coverage of the tablighi jamaat congregation during the COVID-19 pandemic." *Journalism*: 14648849231188260.

Apuke, Oberiri Destiny, Bahiyah Omar, and Elif Asude Tunca. 2023. "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 (1): 88–103.

Ashley, Seth, Adam Maksl, and Stephanie Craft. 2013. "Developing a news media literacy scale." *Journalism & Mass Communication Educator* 68 (1): 7–21.

Badrinathan, Sumitra. 2021. "Educative Interventions to Combat Misinformation: Evidence From a Field Experiment in India." *American Political Science Review* 115 (4): 1325–1341.

Badrinathan, Sumitra, and Simon Chauchard. 2023a. "Researching and countering misinformation in the Global South." *Current Opinion in Psychology*: 101733.

Badrinathan, Sumitra, and Simon Chauchard. 2023b. ""I Don't Think That's True, Bro!" Social Corrections of Misinformation in India." *The International Journal of Press/Politics*: 1–16.

Bandiera, Oriana, Myra Mohnen, Imran Rasul, and Martina Viarengo. 2019. "Nation-building through compulsory schooling during the age of mass migration." *The Economic Journal* 129 (617): 62–109.

Blair, Robert, Jessica Gottlieb, Brendan Nyhan, Laura Paler, Pablo Argote, and Charlene Stainfield. 2023. "Interventions to Counter Misinformation: Lessons from the Global North and Applications to the Global South." *USAID Report*. https://pdf.usaid.gov/pdf$_d$ocs/PA0215JW.pdf.

Bode, Leticia, and Emily K Vraga. 2015. "In Related News, That Was Wrong: The Correction of Misinformation Through Related Stories Functionality in Social Media." *Journal of Communication* 65 (4): 619–638.

Bode, Leticia, and Emily K Vraga. 2018. "See something, say something: Correction of global health misinformation on social media." *Health communication* 33 (9): 1131–1140.

Brashier, Nadia M. 2024. "Fighting Misinformation Among the Most Vulnerable Users." *Current Opinion in Psychology*: 101813.

Brashier, Nadia M, and Daniel L Schacter. 2020. "Aging in an era of fake news." *Current directions in psychological science* 29 (3): 316–323.

Brass, Paul R. 2005. *Language, religion and politics in North India*. Lincoln, NE: iUniverse.

Bridgman, Aengus, Eric Merkley, Peter John Loewen, Taylor Owen, Derek Ruths, Lisa Teichmann, and Oleg Zhilin. 2020. "The causes and consequences of COVID-19 misperceptions: Understanding the role of news and social media." *Harvard Misinformation Review*.

Brulé, Rachel E. 2020. *Women, Power, & Property*. Cambridge University Press.

Cantoni, Davide, Yuyu Chen, David Y Yang, Noam Yuchtman, and Y Jane Zhang. 2017. "Curriculum and ideology." *Journal of political economy* 125 (2): 338–392.

Carpena, Fenella, and Francesca R Jensenius. 2021. "Age of Marriage and Women's Political Engagement: Evidence from India." *The Journal of Politics* 83 (4): 000–000.

Chauchard, Simon, and Kiran Garimella. 2022. "What Circulates on Partisan WhatsApp in India? Insights from an Unusual Dataset." *Forthcoming. Journal of Quantitative Description: Digital Media*. .

Chauchard, Simon, and Sumitra Badrinathan. 2024. "The Religious Roots of Belief in Misinformation: Experimental Evidence from India." *Working paper*.

Clayton, Katherine, Spencer Blair, Jonathan A Busam, Samuel Forstner, John Glance, Guy Green, Anna Kawata, Akhila Kovvuri, Jonathan Martin, Evan Morgan et al. 2019. "Real Solutions for Fake News? Measuring the Effectiveness of General Warnings and Fact-Check Tags in Reducing Belief in False Stories on Social Media." *Political Behavior*: 1–23.

Dahlgaard, Jens Olav. 2018. "Trickle-up political socialization: The causal effect on turnout of parenting a newly enfranchised voter." *American Political Science Review* 112 (3): 698–705.

Davies, Emmerich. 2023. "The Lessons Private Schools Teach: Using a Field Experiment To Understand the Effects of Private Services on Political Behavior." *Comparative Political Studies* 56 (6): 824–861.

Desai, Sonalde et al. 2010. "Human Development in India: Challenges for a Society in Transition.".

Dhar, Diva, Tarun Jain, and Seema Jayachandran. 2022. "Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India." *American economic review* 112 (3): 899–927.

Druckman, James N, and Kjersten R Nelson. 2003. "Framing and deliberation: How citizens' conversations limit elite influence." *American journal of political science* 47 (4): 729–745.

Fazio, Lisa K, David G Rand, and Gordon Pennycook. 2019. "Repetition increases perceived truth equally for plausible and implausible statements." *Psychonomic bulletin & review* 26: 1705–1710.

Flynn, D.J., Brendan Nyhan, and Jason Reifler. 2017. "The Nature and Origins of Misperceptions: Understanding False and Unsupported Beliefs About Politics." *Political Psychology* 38: 127–150.

Gadjanova, Elena, Gabrielle Lynch, and Ghadafi Saibu. 2022. "Misinformation across digital divides: Theory and evidence from northern Ghana." *African Affairs* 121 (483): 161–195.

Gottlieb, Jessica, Claire L Adida, and Richard Moussa. 2022. "Reducing Misinformation in a Polarized Context: Experimental Evidence from Côte d'Ivoire.".

Graham, Matthew H, and Omer Yair. 2023. Less Partisan but No More Competent: Expressive Responding and Fact-Opinion Discernment. Technical report Working Paper.

Guay, Brian, Adam J Berinsky, Gordon Pennycook, and David Rand. 2023. "How to think about whether misinformation interventions work." *Nature Human Behaviour* 7 (8): 1231–1233.

Guess, Andrew, Jonathan Nagler, and Joshua Tucker. 2019. "Less than you think: Prevalence and predictors of fake news dissemination on Facebook." *Science advances* 5 (1): eaau4586.

Guess, Andrew M, and Benjamin A Lyons. 2020. "Misinformation, disinformation, and online propaganda." *Social media and democracy: The state of the field, prospects for reform* 10.

Guess, Andrew M, Michael Lerner, Benjamin Lyons, Jacob M Montgomery, Brendan Nyhan, Jason Reifler, and Neelanjan Sircar. 2020. "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 (27): 15536–15545.

Haas, Nicholas, and Emmy Lindstam. 2023. "My History or Our History? Historical Revisionism and Entitlement to Lead." *American Political Science Review*: 1–25.

Hameleers, Michael. 2020. "Separating truth from lies: comparing the effects of news media literacy interventions and fact-checkers in response to political misinformation in the US and Netherlands." *Information, Communication & Society* 37 (2): 1–17.

Jaffrelot, Christophe. 2021. *Modi's India: Hindu nationalism and the rise of ethnic democracy*. Princeton University Press.

Jha, Himanshu. 2023. "Pathways to develop state capacity in a weak state: the sub-national state of Bihar in India." *Commonwealth & Comparative Politics* 61 (4): 427–450.

Jones-Jang, S Mo, Tara Mortensen, and Jingjing Liu. 2021. "Does media literacy help identification of fake news? Information literacy helps, but other literacies don't." *American behavioral scientist* 65 (2): 371–388.

Kosec, Katrina, and Leonard Wantchekon. 2020. "Can information improve rural governance and service delivery?" *World Development* 125: 104376.

Kumar, Chandan, Piyasa, and Nandita Saikia. 2022. "An update on explaining the rural-urban gap in under-five mortality in India." *BMC public health* 22 (1): 2093.

Kumar, Sanjay. 2019. "In India, Hindu pride boosts pseudoscience." *Science Magazine* 363: 679–680.

Li, Jianing, and Michael W Wagner. 2020. "The value of not knowing: Partisan cue-taking and belief updating of the uninformed, the ambiguous, and the misinformed." *Journal of Communication* 70 (5): 646–669.

Lieberman, Evan S, Daniel N Posner, and Lily L Tsai. 2014. "Does information lead to more active citizenship? Evidence from an education intervention in rural Kenya." *World Development* 60: 69–83.

Malhotra, Pranav, and Katy Pearce. 2022. "Facing falsehoods: strategies for polite misinformation correction." *International Journal of Communication* 16: 22.

Mathew, Athakattu Santhosh, and M Moore. 2011. "State Incapacity by Design." *Institute of Development Studies Working Paper*.

Michelitch, Kristin, and Stephen Utych. 2018. "Electoral Cycle Fluctuations in Partisanship: Global Evidence from Eighty-Six Countries." *The Journal of Politics* 80 (2): 412–427.

Mohan, Janani. 2021. "Media Bias and Democracy in India." *The Stimson Center*. June 28, 2021. https://www.stimson.org/2021/media-bias-and-democracy-in-india/.

Moore, Ryan C, and Jeffrey T Hancock. 2022. "A digital media literacy intervention for older adults improves resilience to fake news." *Scientific reports* 12 (1): 6008.

Mousa, Salma. 2020. "Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq." *Science* 369 (6505): 866–870.

Muralidharan, Karthik, and Nishith Prakash. 2017. "Cycling to school: Increasing secondary school enrollment for girls in India." *American Economic Journal: Applied Economics* 9 (3): 321–350.

Mutz, Diana C, James N Druckman, and Donald P Green. 2021. "Improving experimental treatments in political science." *Advances in experimental political science* 219.

Niemi, Richard G, and M Kent Jennings. 1991. "Issues and inheritance in the formation of party identification." *American Journal of Political Science*: 970–988.

Nyhan, Brendan. 2021. "Why the backfire effect does not explain the durability of political misperceptions." *Proceedings of the National Academy of Sciences* 118 (15).

Paluck, Elizabeth Levy, and Hana Shepherd. 2012. "The salience of social referents: a field experiment on collective norms and harassment behavior in a school social network." *Journal of personality and social psychology* 103 (6): 899.

Pearce, Katy E, and Pranav Malhotra. 2022. "Inaccuracies and Izzat: channel affordances for the consideration of face in misinformation correction." *Journal of Computer-Mediated Communication* 27 (2): zmac004.

Pennycook, Gordon, and David G Rand. 2019. "Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning." *Cognition* 188: 39–50.

Pennycook, Gordon, Ziv Epstein, Mohsen Mosleh, Antonio A Arechar, Dean Eckles, and David G Rand. 2021. "Shifting attention to accuracy can reduce misinformation online." *Nature* 592 (7855): 590–595.

Persily, Nathaniel, Joshua A Tucker, and Joshua Aaron Tucker. 2020. "Social media and democracy: The state of the field, prospects for reform.".

Porter, Ethan, and Thomas J Wood. 2019. *False Alarm: The Truth About Political Mistruths in the Trump Era*. Cambridge University Press.

Porter, Ethan, and Thomas J Wood. 2023. "Factual corrections: Concerns and current evidence." *Current Opinion in Psychology*: 101715.

Prillaman, Soledad Artiz. 2023. *The Patriarchal Political Order: The Making and Unraveling of the Gendered Participation Gap in India*. Cambridge University Press.

Rampal, Nikhil. 2023. "Indian states with less than 10says NITI Aayog report." *National Geographic*. July 18, 2023. https://theprint.in/india/indian-states-with-less-than-10-multidimensional-poverty-doubled-in-5-yrs-says-niti-aayog-report/1673853/.

Rangaswamy, Nimmi, and Supriya Singh. 2009. "Personalizing the shared mobile phone." In *Internationalization, Design and Global Development: Third International Conference, IDGD 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009. Proceedings 3*. Springer pp. 395–403.

Rao, Vijayendra, Kripa Ananthpur, and Kabir Malik. 2017. "The anatomy of failure: An ethnography of a randomized trial to deepen democracy in rural India." *World development* 99: 481–497.

Rasul, Golam, and Eklabya Sharma. 2014. "Understanding the poor economic performance of Bihar and Uttar Pradesh, India: a macro-perspective." *Regional Studies, Regional Science* 1 (1): 221–239.

Roozenbeek, Jon, and Sander van der Linden. 2019. "Fake news game confers psychological resistance against online misinformation." *Palgrave Communications* 5 (1): 1–10.

Roozenbeek, Jon, Sander Van Der Linden, Beth Goldberg, Steve Rathje, and Stephan Lewandowsky. 2022. "Psychological inoculation improves resilience against misinformation on social media." *Science Advances* 8 (34): eabo6254.

Rossini, Patrícia, Jennifer Stromer-Galley, Erica Anita Baptista, and Vanessa Veiga de Oliveira. 2020. "Dysfunctional information sharing on WhatsApp and Facebook: The role of political talk, cross-cutting exposure and social corrections." *New Media & Society*: 14-29.

Scacco, Alexandra, and Shana S Warren. 2018. "Can social contact reduce prejudice and discrimination? Evidence from a field experiment in Nigeria." *American Political Science Review* 112 (3): 654–677.

Sen, Somdeep. 2023. "Big money is choking India's free press — and its democracy." *Al Jazeera*. January 6, 2023. https://www.aljazeera.com/opinions/2023/1/6/big-money-is-choking-indias-free-press.

Sharma, Dinesh C. 2015. "India still struggles with rural doctor shortages." *The Lancet* 386 (10011): 2381–2382.

Siddiqui, Danish. 2020. "Hindu group offers cow urine in a bid to ward off coronavirus." *Reuters*. March 14, 2020. https://www.reuters.com/article/us-health-coronavirus-india-cow-urine-pa/hindu-group-offers-cow-urine-in-a-bid-to-ward-off-coronavirus-idUSKBN2110D5.

Sitrin, Carly. 2020. "New Jersey becomes first state to mandate K-12 students learn information literacy." *Politico*. January 1, 2023 https://www.politico.com/news/2023/01/05/new-jersey-is-the-first-state-to-mandate-k-12-students-learn-information-literacy-00076352.

Taber, Charles S, and Milton Lodge. 2006. "Motivated skepticism in the evaluation of political beliefs." *American journal of political science* 50 (3): 755–769.

Tankard, Margaret E, and Elizabeth Levy Paluck. 2016. "Norm perception as a vehicle for social change." *Social Issues and Policy Review* 10 (1): 181–211.

Tudor, Maya. 2023. "Why India's Democracy Is Dying." *Journal of Democracy* 34 (3): 121–132.

Valeriani, Augusto, and Cristian Vaccari. 2018. "Political talk on mobile instant messaging services: A comparative analysis of Germany, Italy, and the UK." *Information, Communication & Society* 21 (11): 1715–1731.

Vraga, Emily K, Leticia Bode, and Melissa Tully. 2022. "The effects of a news literacy video and real-time corrections to video misinformation related to sunscreen and skin cancer." *Health communication* 37 (13): 1622–1630.