

# Countering Misinformation Early: A Media Literacy Intervention for School Children in India<sup>\*</sup>

Priyadarshi Amar<sup>†</sup>   Sumitra Badrinathan<sup>‡</sup>   Simon Chauchard<sup>§</sup>   Florian Sichart<sup>¶</sup>

## 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 intervention can help school children in Bihar, India, become more discerning consumers of public health information. Collaborating with Bihar’s state government, we developed and fielded an RCT across 583 villages, targeting over 12,000 students, that aimed to engage them 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 6 months after the program ended 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:** Media and Information Literacy, Misinformation, RCT, Education, India

---

<sup>\*</sup>Authors listed alphabetically. This study is funded by SSRC. We are grateful to Jeevika and especially to IAS officer Mr. Rahul Kumar for facilitating the Bihar government’s official collaboration. We also thank Pushpendra Singh Tiwari and Rakesh Kumar at Jeevika. 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, Casey Petroff, and seminar participants at George Washington, Penn, American, Rochester, and UMontreal.

<sup>†</sup>PhD Candidate, University of Wisconsin-Madison

<sup>‡</sup>Assistant Professor, American University

<sup>§</sup>Associate Professor, University Carlos III Madrid & Instituto Carlos 3 - Juan March

<sup>¶</sup>PhD Candidate, Princeton University

# 1 Introduction

Across the world, younger generations exhibit more pro-science attitudes and adeptness at discerning misinformation, highlighting striking disparities between age cohorts (Brashier and Schacter, 2020). For example, survey data underscores higher vaccine hesitancy among older demographics in the US irrespective of partisan leanings (Cao, Ramirez, and Alvarez, 2022). However, Bihar, a state in the eastern region of India with a population of about 127 million, paints a contrasting narrative. Bihar is the poorest state in India: 33% of its population – twice the national average – lives under poverty; a child born in Bihar has a lower likelihood to making it to age 5 compared to a child born in Afghanistan (Kumar, Piyasa, and Saikia, 2022; Rampal, 2023). Bihar also stands as a microcosm of entrenched misinformation. Beliefs in health misinformation attain dramatically high levels among those who stand to lose a lot from harboring them: school children. From our own survey data of 7,000 Bihari adolescents aged 13 to 18, over 87% of respondents said they believe in the curative powers of papaya leaves for dengue fever, while 80% attribute cancer to radiation emitted by cellular towers. Such misinformation may have particularly problematic consequences for individuals who are entering an age period during which core psychological and social developmental processes contribute to making consequential decisions regarding ideology formation, political participation, and health (Persson, Lindgren, and Oskarsson, 2016; Margolis, 2018). Beliefs in medical misinformation like miracle cures can be fatal if even a fraction of those succumbing to them ignore scientific practices (Bridgman et al., 2020).

Such misconceptions, deeply entrenched in cultural and familial legacies, pose a multifaceted challenge: previous work shows that misinformation that taps into traditional cultural belief systems are harder to dislodge (Chauchard and Badrinathan, 2024). This is because the endorsement of misinformed beliefs is also conditioned by familial and social norms, including norms around conformity to group-congruent beliefs and social pressure (Kahan et al., 2017; Tankard and Paluck, 2016). Because the primary locus of childhood socialization in tight-knit village communities is within family units, dislodging misinformed beliefs based on culture and tradition may require individuals to resist pressure from their most direct social networks – families and offline communities.

The prevalence of such falsehoods, however, is not solely a grassroots phenomenon confined to rural communities. Its dissemination is also top-down, as misinformation often originates from or resonates within the political arena, where identity politics and nationalist fervor coalesce. In India, the entanglement of health remedies with cultural and religious symbolism renders misinformation a potent political tool wielded by influential Hindu-nationalist politicians, particularly elites aligned with the Bharatiya Janata Party (BJP). In March 2020 a BJP group in New Delhi, India organized a 200-person event to encourage drinking cow urine to cure the virus (Siddiqui, 2020), blurring the line between traditional misinformed beliefs and political agendas. In 2021, some state governments handed out herbal medicine to those waiting in line to get vaccinated. Compounding these challenges are systemic deficiencies in state capacity and information dissemination. Traditional media, increasingly aligned with governmental narratives, amplifies partisan voices while diminishing dissent (Mohan, 2021; Sen, 2023). Further, in rural Bihar where healthcare infrastructure is weak, quack doctors proliferate, further exacerbating health disparities (Sharma, 2015). This nexus of elite-backed disinformation, limited access to credible news, and institutional frailties creates an ecosystem where misinformation thrives with impunity. Families, schools, and the state — all foundational pillars of societal edification — falter in their roles, leaving adolescents facing an information void, ripe for exploitation.

How can we effectively address misinformation in this context? To answer this question, we deploy a long-term media literacy education intervention across 583 villages in Bihar, targeting over 12,000 adolescents aged 13-18.

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.<sup>1</sup>

In this study, we attempt to address this gap by exploring this less conventional and understudied approach to counter misinformation, namely, sustained media and information literacy education 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.

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 Jayachandran, 2022; Badrinathan, 2021). Topics explored included the allure of emotional content, the influence of confirmation bias and motivated reasoning on vulnerability, techniques for engaging in respectful and corrective dialogue with individuals espousing falsehoods, and avenues for verifying information credibility. The program’s messaging aimed to raise awareness about the detrimental effects of misinformation while furnishing practical methods for enhancing information processing skills in an difficult informational landscape. The design of the treatment curriculum was informed by insights gleaned from consultations with subject mat-

---

<sup>1</sup>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.

ter experts, local educators, journalists, and a review of academic literature delineating effective strategies for combating misinformation. DataLeads, a non-profit organization based in Delhi, India, operationalized the curriculum into actionable lesson plans, including slide decks, flyers, and graphics. Additionally, DataLeads recruited and trained educators to deliver lessons. 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 evaluate the effect of this intervention on a number of primary misinformation outcomes, including both beliefs and incentivized 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 stories relying on traditional beliefs and religious systems, is more difficult to change (Chauchard and Badrinathan, 2024). Further, the treatment enhanced assessment of sources and their ability of treated students to discern their credibility. However, the intervention did not improve self-reported understanding of the underlying theoretical concepts around misinformation, nor did it alter engagement behaviors with misinformation countermeasures.

What is especially striking is that these effects persisted over time and we detect spillover effects to students networks. We resurveyed a random sub-sample of 2,000 participants 6 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 new domains. Finally, we also find that parents of treated students are better able to discern true from false information.

The implications of these findings hold significant importance, especially in light of prevailing challenges faced by most media 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 communal 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 resembling 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 Motivating Media Literacy

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 in the Global South (Blair et al., 2023; Badrinathan and Chauchard, 2023a).

Our intervention consists of a classroom-based media literacy program, entitled the Bihar Information and Media Literacy Initiative (BIMLI). Media literacy is broadly defined as the



skills and competencies that promote critical engagement with messages produced by the media, needed to successfully navigate a complex information ecosystem (Jones-Jang, Mortensen, and Liu, 2021). In practice, however, media and information literacy 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. Our approach relied on combining several of the best practices and definitions outlined in previous research, and putting together these insights into a coherent curriculum. 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.

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. We chose an offline, classroom-based intervention, thereby departing from online interventions relatively popular in misinformation studies (see for example Roozenbeek et al. (2022); Pennycook and Rand (2019); Blair et al. (2023)) as only a small proportion of our sample has first-hand access to the online world. As is confirmed by our baseline data, less than 1 in 5 of our sample reported using a cellphone. Moreover, in those cases where households did have internet access, our field enumerators noted this was largely through the use of mobile phones shared by all or several members within a household, marking a divergence from Western settings (Rangaswamy and Singh, 2009). Further, when personalized access to the internet exists in India, this takes place largely through the medium of WhatsApp, the most popular platform in the country. WhatsApp, characterized by encrypted peer to peer sharing, makes online alternatives to our intervention such as such as corrections and fact-checking, or algorithmic changes, inappropriate and unfeasible (Badri-nathan, 2021; Rossini et al., 2020; Valeriani and Vaccari, 2018).<sup>2</sup> Consequently, we opted for a sustained, iterative approach involving multiple sessions of learning as well as homework as-

---

<sup>2</sup>While scholars have looked at social corrections or peer corrections to misinformation on encrypted apps (Badri-nathan 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.



Table 1: Description and Objectives of Media Literacy Curriculum

	Learning Objectives	Theoretical Framework	Indian Context
Module 1	<u>Definitions</u> : 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	<a href="#">Guess and Lyons (2020)</a> '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	<u>Causes of vulnerability</u> : human psychological biases like confirmation bias; news and media system biases such as financial incentives; the need for critical thinking	Motivated reasoning from <a href="#">Taber and Lodge (2006)</a> , list of psychological biases adapted from <a href="#">Roozenbeek et al. (2022)</a> , list of media biases adapted from <a href="#">Ashley, Maksl, and Craft (2013)</a> , news framing effects from <a href="#">Druckman and Nelson (2003)</a> , fact-opinion discernment from <a href="#">Graham and Yair (2023)</a>	Introduction to the news media environment in India + how biases manifest in the Indian context (for example, scapegoating minorities)
Module 3	<u>Solutions to misinformation</u> : 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 ( <a href="#">Guess, Nagler, and Tucker, 2019</a> ; <a href="#">Vraga, Bode, and Tully, 2022</a> ; <a href="#">Badri-nathan, 2021</a> ), focus on sharing as different from belief ( <a href="#">Brashier and Schacter, 2020</a> )	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	<u>Family, vaccination</u> : 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 ( <a href="#">Bode and Vraga, 2018</a> ; <a href="#">Badri-nathan and Chauchard, 2023b</a> ), talking to older family about misinformation ( <a href="#">Pearce and Malhotra, 2022</a> )	Role-playing exercise to talk to an Indian relative about misinformation

signments between sessions, thereby departing from almost all other educative interventions in the misinformation literature (Blair et al., 2023).

Thus BIMLI is, by design, a bundled intervention: it aimed to cultivate awareness of misinformation, particularly within the realm of health. It was also designed to foster a reliance on and trust in scientific information, and to equip participants with practical skills and strategies to combat misinformation in real-world scenarios. This included understanding human psychological biases, discerning the origins and dissemination of misinformation, and acquiring techniques to identify and trust credible sources. Finally, BIMLI seeks to provide participants with an opportunity to navigate the challenges inherent in a complex informational landscape characterized by a dearth of reliable information, as well as to empower them to engage in dialogue and challenge individuals, particularly elders and those around them susceptible to misinformation. 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.

The total dosage of our program amounted to approximately 10 contact hours between

facilitators and students. While this duration may appear modest compared to other RCTs within traditional classroom settings, we underscore that the frequency and intensity of BIMLI was far stronger in dosage in comparison to existing media literacy interventions in the academic literature (Blair et al., 2023). Several studies use media, information, and news literacy campaigns, and indeed, many of their treatments are much less laborious. For example, Hameleers (2020) and Guess et al. (2020) show respondents tips and strategies to spot misinformation via flyers, Ali and Qazi (2021) show respondents a 3-4 minute media literacy video. Closer in principal to the current study, Badrinathan (2021) discusses misinformation and concepts related to media literacy with respondents in an hour-long intervention, and Moore and Hancock (2022) provide respondents with an hour-long self-directed series of modules on digital literacy. Crucially, most existing studies are one-off treatments with no sustained learning component; besides, in many cases, self-described media or digital literacy treatments simply provide respondents with primes or reminders about critical thinking, rather than actual tools or techniques for learning. More importantly, most existing media literacy initiatives in the global south have mixed to null results (Blair et al., 2023). Consequently, we designed a strong dosage treatment with the aim of inducing significant variation in the independent variable (Mutz, Druckman, and Green, 2021), in a context and population where the effect of such interventions remains untested.

### 3 Study Design and Data Collection

We implement a randomized control trial to test the efficacy of the BIMLI media literacy 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).

#### 3.1 Administering classes

To bolster the credibility of BIMLI, we secured official collaboration with an agency of the Bihar local government, the Bihar Rural Livelihoods Promotion Society (BRLPS, or as it is commonly known, Jeevika), to launch BIMLI statewide. Jeevika has a history of collaborating with academic

institutions to implement RCTs throughout Bihar. 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 and alternative medicine sources among youth, and especially in rural regions.

In our study, participants were students enrolled 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 covered multiple classrooms across 2-3 districts. The facilitator visited each classroom roughly once every two to three weeks. To hire teachers, we collaborated with DataLeads, an organization based in Delhi, India, that is an official partner for the Google News Initiative and has years of experience in the media literacy and fact-checking space in India.<sup>3</sup>

To dispense the classes, Jeevika made available to us 99 community libraries across 32 districts (one in each block). We ran our classes in these libraries. We purposefully 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 very overburdened, that absenteeism (in students and teachers) was very common in Bihar, and that running the intervention in schools may increase the risk of spillover.<sup>4</sup> 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

---

<sup>3</sup>DataLeads put out a call to hire trainers in Bihar to deliver our intervention; out of the 400 applicants who applied to the call, DataLeads selected a final pool of around 50 teachers based on an interview process, and then conducted a 2-day training where they assessed the skills and interests of those who were shortlisted. One of the criteria for hiring was to have a dynamic teaching style to ensure a discussion-based classroom style so as to maintain student participation and interest. The final pool of selected teachers comprised not only of existing school teachers and educators, but also journalists, college professors, and some fact-checkers. Everyone selected was in-state, with most teachers tasked with teaching classes in their own sub-region or district of Bihar so as to minimize travel time. Teacher schedules were drawn up such that each classroom of students had the same teacher throughout the intervention, but each teacher was assigned 6-9 batches or classrooms on average, across 2-3 districts.

<sup>4</sup>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.

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, Jeevika provided access to government-owned libraries established in each district throughout the state, to be used as classrooms for the program. These newly established facilities not only offered familiar environments for students and their parents but also provided 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. Moreover, we also expected that public schools would have a lower appeal for children and that they might be more motivated to attend in a slightly different setting.<sup>5</sup> Figure 1 provides an example of a BIMLI classroom in session.

Figure 1: BIMLI trainer explains a role-playing exercise to students in Vaishali, Bihar.



Collectively, these efforts served to incentivize enrollment in the program and ensure

---

<sup>5</sup>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 of every one of the 28 states in India that the survey spans.

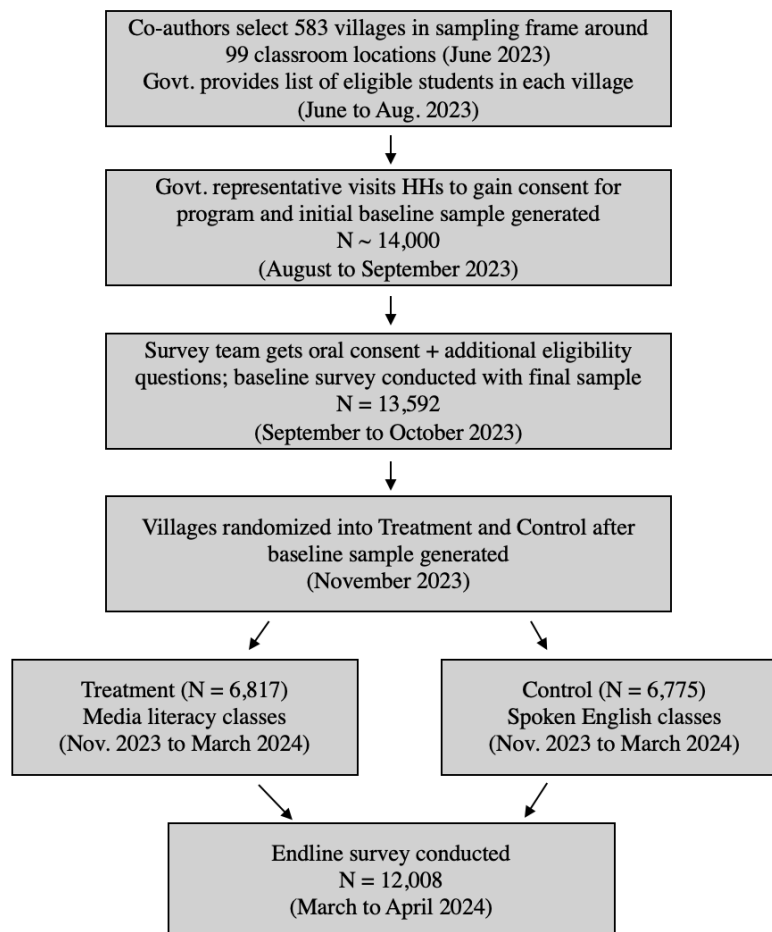
sustained participation throughout its duration. Incentivizing students to attend by increasing program legitimacy was one of our central goals while designing this study. While educational interventions have normative value, we highlight that past studies using such methods, especially in the global south, have often been met with null results (Badrinathan, 2021; Ali and Qazi, 2021; Blair et al., 2023). This ties into a broader literature in comparative politics highlighting the frequently null effects of information provision, implying that information may be insufficient to elicit changes in belief and behavior. Many citizens are constrained by factors that impede the translation of new information into actionable responses. Central among these factors is the notion that citizens must not only be exposed to novel information but also incentivized to prioritize the issue and believe in the efficacy of their actions (Kosec and Wantchekon, 2020). Additionally, scholars emphasize the potential influence of group dynamics, wherein the knowledge that others are similarly informed may exert pressure and facilitate changes (Lieberman, Posner, and Tsai, 2014). Consequently in our study context, the educational classroom setting and the iterative nature of the intervention, characterized by exposure to authoritative figures and peer interactions, holds promise for fostering enhanced receptivity to information interventions.

In addition, null effects may stem from a lack of support from governing authorities as well as disparities in implementation facilitation (Rao, Ananthpur, and Malik, 2017). Particularly in developing countries, the absence of elite involvement in intervention rollouts may hinder citizen mobilization, especially in rural areas. Furthermore, the efficacy of participatory interventions is contingent upon proficient facilitation and consistent implementation standards. Our study incorporated key design choices to accommodate these considerations. These included formal collaboration with Jeevika to facilitate outreach in rural areas, thereby ensuring broader and sustained citizen engagement. Government-endorsed initiatives, particularly those addressing misinformation, serve to shield citizens from potential elite resistance while providing access to pertinent information on sensitive topics (Rao, Ananthpur, and Malik, 2017). Additionally, to ensure effective facilitation, external educators were enlisted, interventions were conducted in adequately resourced library settings, and a comprehensive field monitoring plan was instituted to oversee compliance and mitigate attrition among both educators and students (we describe monitoring and compliance in greater detail in sections below).

### 3.2 Sampling, enrolment and baseline data

Figure 2 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 2: Flow, Timeline, Enrollment and Randomization

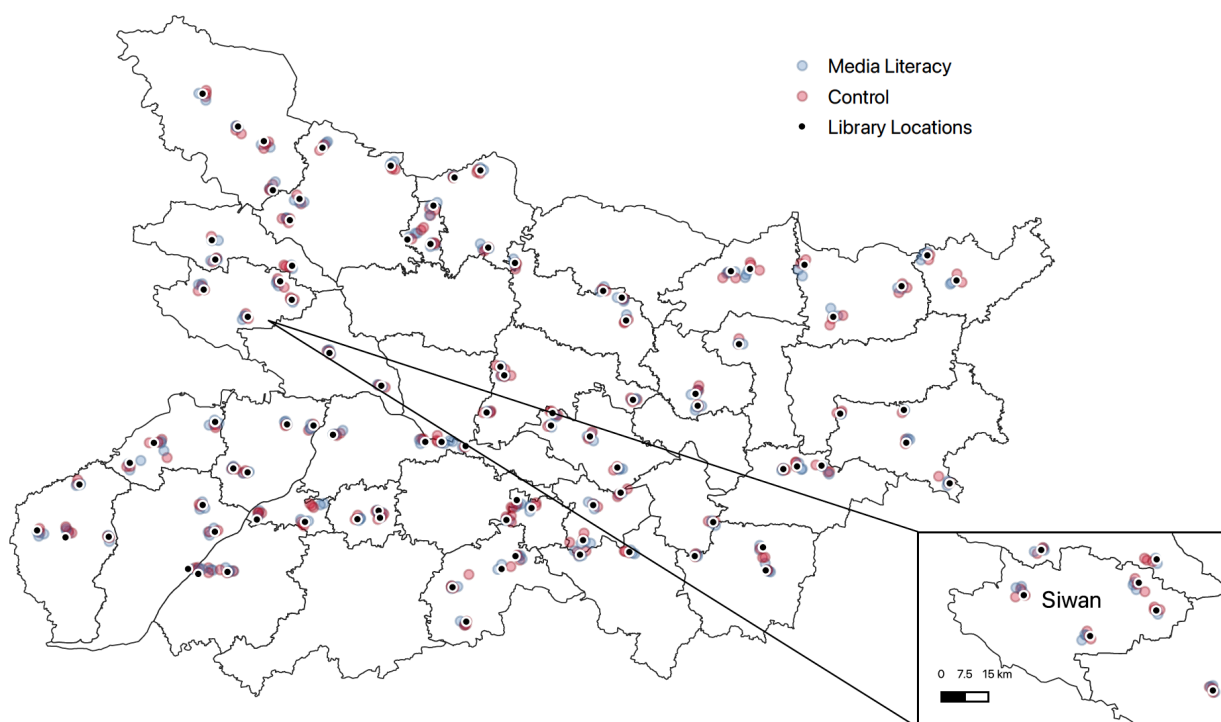


Within each of the 583 selected villages, we then relied on Jeevika to provide a list of students eligible for the study, based on existing household data that the government has from enrolment in state programs. Initial criteria for eligibility in the study included households with children enrolled in government schools in grades 8 to 12. Once a long list of such households



was generated, a representative from Jeevika visited these households to first confirm that an age-eligible and school-going child was present. If this was the case, Jeevika staff proceeded to provide a short description of the program and asked the student as well as a parent or guardian present if they were interested in participating. If students and parents agreed after this initial pitch, a shorter list of (20 to 24) eligible and interested students was generated within each selected village. Figure 3 shows a map of the treatment and control villages in Bihar along with library locations.

Figure 3: Map of treatment and control villages



Crucially, we note that randomization occurs *after* students opt in, so that the issue of differential opt-in rates between treatment and control is avoided. On this point, everyone involved in this study – including teachers, our implementation partners, government officials, and coauthors – were blind to treatment status at the time students were recruited and during baseline data collection. During the initial household visit, students and their parents were merely told that they would be given the free opportunity to go through a government-offered certificate course with 4 sessions, the content of which has benefits for their future and careers as they go

off to college or on the job market.

Next, our survey team visited eligible and interested students to conduct the baseline survey. The baseline survey was conducted by Sunai Consultancy, Patna, between September to October 2023. Enumerators visited each house on the Jeevika-generated list in-person for the baseline survey, which included three additional eligibility criteria. First, enumerators obtained official oral consent from children and one parent or guardian present. Second, we included a one-item measure of students' basic (second-grade) reading comprehension in Hindi. If students failed or performed badly on this item, the household in question was replaced in our sampling frame.<sup>6</sup> Third, we asked both students and parents to verbally affirm that the student would be able to attend four classroom sessions during the stipulated weeks and timings and at their library location; if either said no we dropped the household from the sample and replaced it.

After students and their parents opted in, the baseline survey was completed and the final sample was generated, we randomized and assigned half of the included villages to treatment based on the criteria described in Appendix A. As such, students did not know whether they were receiving media literacy classes or control until they showed up to the first day of the sessions.

The baseline survey measured basic demographics and basic household level data, along with attitudinal modules including perceptions of the state and institutions, media access and usage, opinions on science and vaccines, social ties, and a personality inventory. Our baseline sample counts  $N = 13,592$  respondents across 583 villages, with 49.9% of the sample assigned to treatment and 50.1% assigned to control. In our sample, 58% of respondents are female. Students range from grades 8 to 12, with the median class being Grade 10. 96% are enrolled in government schools. Further, the sample is 91% Hindu and 69% OBC. Religious and caste variables line up with the overall demographics of the state as reported in the census. We count diversity in language spoken at home, with 43% from Hindi-speaking households, 30% Bhojpuri and about 9% Magahi. On average students come from households in which fathers' median education

---

<sup>6</sup>This additional eligibility criteria was included based on existing data about poor reading and writing levels of public school children in rural Bihar. ASER's 2022 report demonstrated that over 30% of school children 13 years and over were unable to read a 2nd grade level text in Hindi. Though our classes or endline did not include a reading or writing component, we determined it necessary to have basic language comprehension to fully grasp the content of the intervention.

Table 2: Descriptive Statistics: Student and HH Characteristics at Baseline

Variable	Treatment %	Control %	p-value of diff.
Male	42.23	41.33	0.58
Hindu	90.98	91.44	0.40
School Grade (mean)	9.62	9.70	0.017
Government School	96.62	96.64	0.72
General Caste	8.5	7.6	0.82
OBC	69.66	68.29	0.87
Hindi	44.1	42.41	0.89
Bhojpuri	30.96	30.37	0.85
Family eats meat	90.34	90.51	0.70
Father's education (mean years)	6.9	6.9	0.92
Mother's education (mean years)	4.2	4.1	0.65
HH owns washing machine	3.51	3.76	0.93
HH owns fridge	15.5	16.01	0.77
HH uses mobile internet	19.48	18.92	0.93
Vaccinated	76.22	78.19	0.20
Believes Ayurveda effective	71.9	71.61	0.91
Science knowledge scale (mean correct)	4.55	4.49	0.80
<i>N</i>	6,775	6,817	

Note: p-values from chi-squared tests with [Rao and Scott \(1987\)](#) second-order correction.

was grade 6-9 and mothers' median education was grade 1-5. In addition, we measured the presence of a number of household assets to proxy for socio-economic status. Of note, 15% of the sample came from households owning a refrigerator and only 3.6% from households owning a washing machine. 19% of the sample reported having an active mobile phone with internet. On attitudinal variables, respondents reported a high degree of trust in the media; 90% said they trusted newspapers, 84% television and 61% information on social media. A majority of the sample (77%) reported having been vaccinated against covid-19. Despite this, 87% said they believed alternative medicinal techniques such as ayurveda and homeopathy were effective.

Table 2 reports baseline characteristics of the sample by treatment status. The data confirms that the two samples are balanced on demographic characteristics as well as behavioral and attitudinal characteristics that might predict dependent variable measures, such as vaccination rate, science knowledge, or internet access. We do detect a statistically significant difference between the average grade of students in the control group (9.70) versus students in the treatment group (9.62), though this difference is not substantively large.

### 3.3 Control condition

Control group units received four modules of classroom-based in-person conversational English language classes. Our goal in doing so was primarily to have a placebo, rather than pure control group, so as to be able to achieve parity between the treatment and control groups in terms of exerted effort among students. Acknowledging the pivotal role of sustained compliance and active engagement in the efficacy of iterative classroom interventions, we strived to generate comparable classroom dynamics and levels of peer interaction, introducing variation solely in the content of instruction. Additionally, Jeevika encouraged us to ensure that even control group participants gained practical benefit from the program. Consequently, we considered a number of pedagogical themes for control group classes, with the dual goal of fostering engagement while mitigating potential confounding influences on misinformation outcomes. Notably, we excluded subjects such as mathematics due to their perceived redundancy with standard school curriculum, as well as the lack of enthusiasm students had for this subject during preliminary fieldwork. Besides, science disciplines were deemed unsuitable because of their obvious associa-

tion with the treatment topic (part of the intervention emphasizes types of evidence, scientific or not, for instance), whereas history and geography were unsuitable due to their link with national identity narratives. We considered alternative non-textbook domains, including culinary arts, handicrafts, and other hands-on activities, but ultimately discarded these due to the foreseeable gendered nature of their uptake, and because we thought students stood to benefit less from them.

Ultimately, we decided to implement a curriculum consisting of four sessions focusing on basic conversational English. The primary objective of these sessions was to equip students with foundational language skills aimed at enhancing their prospects in future career endeavors and job interviews. Notably, these modules deliberately omitted written and reading components, with the instruction solely concentrated on spoken phrases and sentences. We highlight that the instructional content was tailored to cater to students with limited or no prior exposure to the English language, particularly those primarily educated in non-English medium schools. Consequently, the modules were crafted at a level akin to that of first-time learners.

The curriculum started by elucidating to students the significance of being proficient in expressing themselves in English within the context of career opportunities. Subsequent modules encompassed topics such as self-introduction, naming common objects, describing everyday activities, and formulating questions. A detailed overview of the lesson plans is provided in Appendix C. We note that we dedicated attention to aligning the structure and flow of these classes with those of the treatment group. Activities such as role-playing, group exercises, and homework assignments prompting discussions with parents were incorporated in a manner mirroring that of the treatment sessions. These sessions deliberately refrained from referencing topics such as media, technology, or politics. Due to the very basic nature of this instruction, students are unlikely to have acquired sufficient proficiency in English, or any other discernible skills, by the conclusion of the course to navigate new information sources or the internet independently.

We note here that the teacher selection process and training differed for treatment and control groups. While teachers in the treatment group were recruited and trained by DataLeads, teachers for the English class were recruited and trained by a local Bihar consultant. Because our sample of teachers did not draw from the same overall pool, they have different socio-economic characteristics and experiences with teaching. Given these potential differences, the treatment

effects we measure are, by design, a function of the content of treatment and of the characteristics of the two pools of teachers. In Appendix G we summarize demographic characteristics of teachers by treatment condition.

### 3.4 Endline data and compliance

Our first endline survey was conducted after 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 logistical constraints, 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 6 months after the intervention, to assess if treatment effects persisted. This survey was conducted over the phone with a random subset of 2,000 students.

To ensure compliance with classes among both students and teachers, we instituted a meticulous monitoring procedure. First, concerning students, Jeevika effectively engaged their local staff members, referred to as *didis*, to periodically remind students and households of impending classes. While there was variation in the motivation levels of these *didis* at the local level, given their status as bureaucratic actors constrained by time and resources (Dasgupta and Kapur, 2020), they were repeatedly requested by Jeevika hierarchy throughout the implementation phase to facilitate student adherence to the program. In addition we also provided monetary compensation to each *didi* for their services throughout the study period. Further, students were

independently motivated to attend classes due to our promotion of the program as a government-certified course. Upon completion of the four classes, attending students received a certificate in an official ceremony conducted by a Jeevika representative. Moreover, during class sessions, students were provided with various materials to take home, including notebooks, pens, informational flyers, and homework assignments, alongside snacks at the conclusion of each session.

Table 3: Attendance for full sample

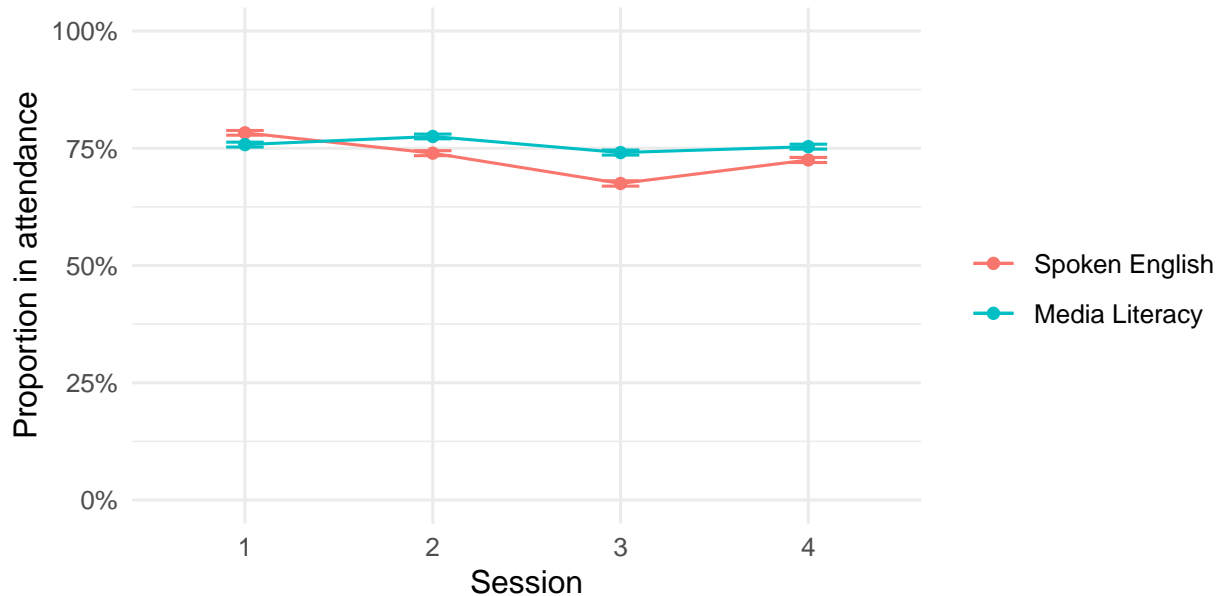
	Percent
Attended 0 classes	10.02
Attended 1 class	7.70
Attended 2 classes	9.73
Attended 3 classes	19.84
Attended all 4 classes	52.69

Finally, on the teacher side, we recognize that despite training and remuneration, teachers may be unable to show up to classes due to multiple context-specific barriers such as commuting distance or unforeseen circumstances that may hinder their presence in classes. Given the extensive scope of our study, we had external monitors conduct random visits to village libraries to verify the adherence to the stipulated schedule and the presence of teachers. Subsequent to each visit, these external field monitors completed a concise survey documenting their observations from the field. In addition, during the initial and final stages of classroom sessions, two of the co-authors alternated in visiting these classes to oversee their smooth operation. Overall, this leads us to have a high degree of confidence in both our attendance data and in the standardized nature of the intervention. Moreover, throughout the study, teachers were asked to upload through an app materials documenting attendance and documenting that the session had indeed taken place on that day, among which featured photos of the board they had used. To enter attendance, teachers filled out an attendance sheet at the end of every session, which included respondent-level attendance in the classroom for each session.

In Table 3 we tabulate the percent of students who attended 0 through 4 sessions of the course. 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



Figure 4: Compliance data across treatment and control



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 4 graphs attendance data by treatment group and by session.

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, or trust in the media.

### 3.5 Outcome measures

We pre-specified seven primary families of outcomes for the endline survey: awareness of misinformation, accuracy discernment, sharing discernment, health attitudes, trust in sources, demand for fact-checking, and engagement with misinformation efforts.<sup>7</sup> 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.<sup>8</sup> The specific variables and procedure for constructing the index were pre-specified (see Appendix F). We describe our primary outcomes as we present our main results, in the next section.<sup>9</sup>

## 4 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:

$$Y_{ijk} = \beta_0 + \beta_1 T_{ij} + \sum_{k=1}^{m-1} \gamma_k + \varepsilon_i \quad (4.1)$$

$$Y_{ijk} = \beta_0 + \beta_1 T_{ij} + \sum_c \alpha_c X_{ci} + \sum_{k=1}^{m-1} \gamma_k + \varepsilon_i \quad (4.2)$$

where  $Y_{ijk}$  is the primary outcome of interest  $Y$  for student  $i$  in classroom  $j$  and district  $k \in \{1, \dots, m\}$ ,  $\beta_0$  is the intercept,  $T_{ij}$  is a treatment indicator,  $\alpha_c$  denotes the coefficient for the control variable  $X_c$ <sup>10</sup>,  $\gamma_k$  denotes fixed effects for each district  $k$ , and  $\varepsilon_i$  denotes the random error term

<sup>7</sup>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>.

<sup>8</sup>For details on the rationale behind using ICW indices and how exactly we construct the outcome measure, see Appendix F.

<sup>9</sup>In addition to these primary outcomes, we pre-registered examining several secondary outcomes to ascertain whether the intervention yields spillover effects within families. This entails assessing whether children discuss program content at home, thereby potentially enhancing the discernment abilities of parents in distinguishing between true and false information. We will also investigate whether the intervention induces shifts in attitudes towards institutions, including perceptions of schooling, trust in authorities, and civic values. Finally, we pre-registered heterogeneous treatment effects hypotheses regarding the efficacy of the treatment on (a) those with low/high prior access and use of media and (b) those with lower prior attachment to non-scientific belief systems.

<sup>10</sup>This model will include the following baseline controls: individual level (age, gender, medium of education, grade

for individual  $i$ .  $\beta_1$  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.

We examine the overall effect of BIMLI by comparing endline outcomes for those in the media literacy 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 5 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, improving discernment of true from false information, discernment in sharing intentions, discernment between sources of news, health preferences, and attitudes towards engaging with fact-checking. Despite these changes, we do not detect an average effect of awareness about the different facets of misinformation or willingness to engage with misinformation counter-measures. We detail these results in the following subsections.

## 4.1 Main effects

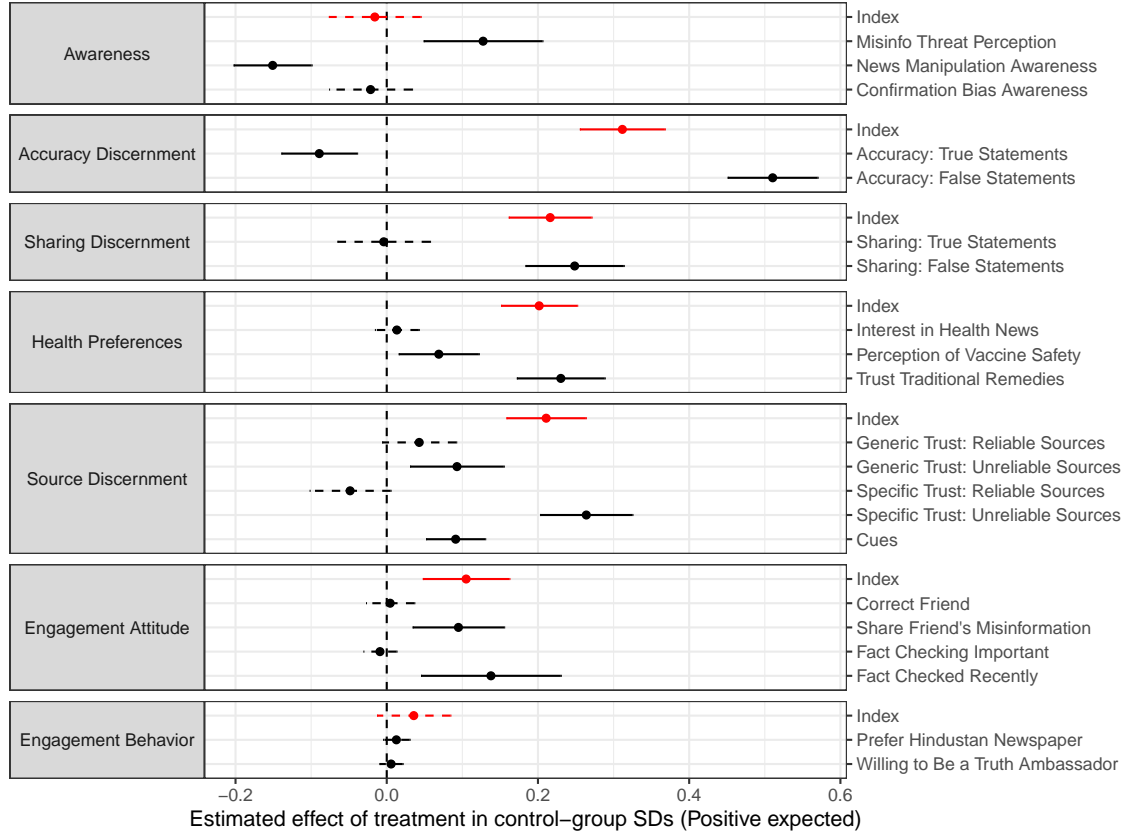
### 4.1.1 Awareness of misinformation

To gauge awareness of misinformation as a societal issue as well as awareness of the material of classroom sessions, we constructed an ICW index comprising of five items (detailed in Appendix F). The first item evaluates respondents' perception of misinformation as a threat, measured on a scale of 1 to 5, juxtaposed with other societal issues such as unemployment or corruption. Although we did not observe a discernible effect on the overall awareness index, we found that exposure to BIMLI significantly increased respondents' perception of misinformation as a threat. It is worth noting, however, that a substantial proportion of respondents (78%) initially perceived misinformation as a considerable threat to begin with (scoring 4 or 5). Consequently, our capacity to influence respondents on this metric might have been constrained.

---

eg. class 9-12, reading and science skill index, prior exposure to internet), household (assets, religion, caste category), and village level (development, proxied by night lights data, BJP vote share at the assembly constituency level)

Figure 5: 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 I.

Next, we measure awareness of biases in the way news is produced as well as awareness of cognitive/psychological biases amongst consumers of news through a series of items adapted to the Indian context from (Ashley, Maks, and Craft, 2013). We include 2 items each on media biases and cognitive biases, scored on a strongly agree to strongly disagree scale. An example of a cognitive bias item is: "People pay more attention to information that fits their beliefs than to information that does not". Regarding cognitive bias items, treated respondents exhibited no discernible improvement relative to the control group, and we could not reject the null hypothesis of zero effect ( $p = 0.7$ ). It is important to note that these items primarily tested students' ability to recall concepts taught in class, such as the definition of confirmation bias, rather than the application of classroom-acquired knowledge. Hence, it is plausible that respondents found them comparatively challenging to answer, especially given the time gap between

the session that discussed these concepts and the endline survey. Notably, the observed effects on discernment metrics (discussed below), juxtaposed with the absence of significant effects on these awareness items, suggest respondents' capability to leverage classroom instruction to inform their information consumption practices, notwithstanding potential difficulties in recalling theoretical definitions and concepts from lectures.

Conversely, concerning news manipulation items, treated respondents demonstrated a reduced propensity to identify instances of bias in the media. An example of such an item is: "News organizations select stories based on their potential to attract the largest audience." Notably, respondents in the treatment group were more inclined to disagree than agree, counter to expectations. These counter-intuitive outcomes could stem from the phrasing of the questions: qualitative insights underscore potential respondent confusion in distinguishing between their perceptions of what news companies ought to do and their observations of what they actually undertake. Consequently, while we observed an overall null effect on the index, there was significant variation in the way respondents answered each component.

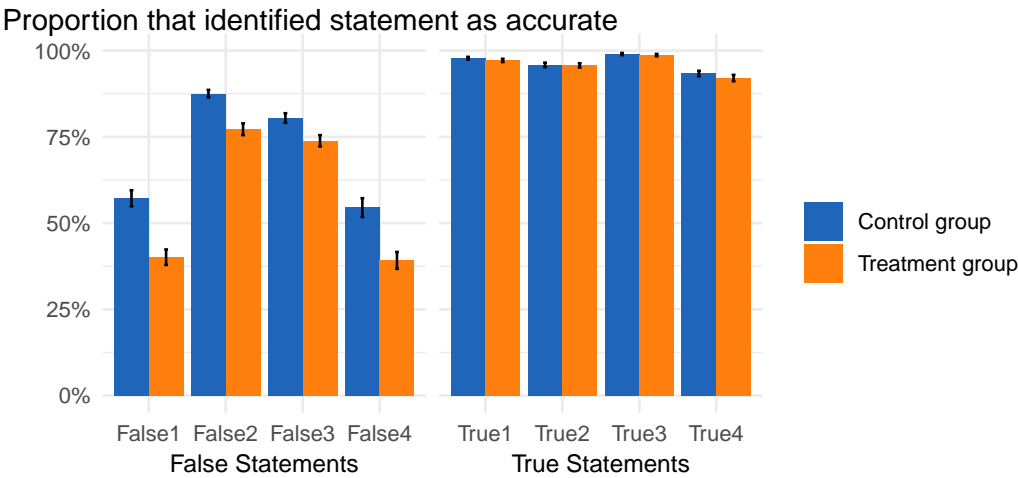
#### **4.1.2 Accuracy and sharing discernment**

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 primarily a function of learning and application, rather than mere recall. We used the same set of 8 stories to measure sharing

discernment by 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 the stories utilized this measure was guided by several key considerations. Primarily, we aimed to incorporate stories that respondents were already familiar with and which held salience. This served a dual purpose: firstly, to avoid introducing unfamiliar misinformation narratives, and secondly, to ensure that the selected stories were prevalent within the cultural landscape, thus allowing for the treatment to potentially influence beliefs. Our final selection of stories stemmed from extensive fieldwork and piloting efforts, aimed at identifying the most commonly believed health-related misinformation stories within our context. Further, we took care to ensure that these narratives were not merely community-based rumors but were also news headlines. Each false story ultimately chosen had been independently verified as false by at least one third-party fact-checking service in India. We presented the headlines to respondents in random order.

Figure 6: 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 5). The magnitude of this effect, a 0.32 SD increase in discernment relative to the control group, is substantively large compared to other work in both similar and dissimilar contexts. 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 6](#) 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. Thus while the treatment marginally made respondents more skeptical of true stories, these effects are substantively very small.

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). When broken down into true and false components, the treatment significantly reduced willingness to share false stories (0.25 SD) and did not have an effect on true stories (0.00 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, 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.<sup>11</sup>

---

<sup>11</sup>Tabular results for these analyses are available in Appendix I.



### 4.1.3 Health preferences

One of the principal objectives of BIMLI was to facilitate a shift in preferences regarding health-related behaviors. Our fieldwork revealed the prevalence of misinformation pertaining to non-scientific and alternative forms of medicine. Jeevika's support for our RCT also stemmed from the government's aim to address and potentially modify individuals' (misinformed) beliefs regarding health preferences.

In our endline we measured health preferences via three components: interest in health news, perceptions of vaccine safety, and reliance on traditional/alternative medicine. To measure interest in health news we asked respondents how interested they were on scale of very interested to not interested in health among other news topics. We employed two items to measure perceptions of vaccine safety. These ask, on a scale of safe to unsafe, how respondents feel about the safety of the covid-19 vaccine and the chicken pox vaccine. Importantly, to measure reliance on alternative medicine, we ask respondents whether they would engage in the following activities to respond to serious medical illnesses in their household: (a) going to a non-scientific traditional healer, (b) treating it at home with home remedies. We also ask on a 4-point agree disagree scale the extent to which respondents agree that ayurveda and homeopathy have the capacity to cure serious illnesses.

We found significant variation in responses to these questions at baseline. For instance, over 77% of respondents at baseline reported being vaccinated against Covid-19, with an even larger proportion expressing intent to receive vaccination. This inclination may reflect trust in the government, as vaccination campaigns were actively promoted by the central government through a coordinated nationwide rollout. Concurrently, 87% of respondents at baseline asserted that alternative forms of medicine, such as ayurveda and homeopathy, were effective means of treating illnesses. Many participants in our fieldwork discussions mentioned relying on health advice from friends and family rather than trusted sources like doctors or healthcare professionals. Additionally, several respondents highlighted the prevalence of home remedies and medical misinformation within their information networks, 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 did not influence interest in health news, it 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.<sup>12</sup>

#### 4.1.4 Trust in sources and source discernment

We introduced a series of novel items to gauge how respondents navigate trusting various sources and discerning between sources of news and information. In the literature, the standard approach to measuring trust in mainstream news involves classifying various news brands (such as CNN or Fox News) as trustworthy or untrustworthy. However, we opted against this approach for two main reasons. First, such a classification would be challenging to justify within the context of India because of recent political developments eroding the quality of independent media.<sup>13</sup> Second, we were operating within a context where respondents heavily relied on informal sources of news, such as in-person conversations or word of mouth, as well as on local elites

---

<sup>12</sup>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.

<sup>13</sup>Discourse analyses on Indian mainstream media outlets suggest that it is hard to distinguish between right versus left and trustworthy versus untrustworthy sources, as all mainstream brands have engaged in hateful rhetoric, including misinformation, exaggerations, and polarizing content over the last decade (Amanullah, Nadaf, and Neyazi, 2023; Mohan, 2021).

(Gadjanova, Lynch, and Saibu, 2022). Considering these factors, our index of source discernment encompasses three components.

First, we measure general source discernment by asking respondents to rate, on a scale strongly trust to strongly distrust, their level of trust in a variety of sources. We conceptualize “sources” broadly, including transmitters of information (such as word of mouth), mediums through which information is disseminated and received (such as radio or Facebook), and institutions (such as the WHO or central government announcements). For our index, however, we focus on four distinct sources: two of which we anticipate our treatment will increase trust in (MBBS doctors and healthcare workers such as ASHA), and two of which we hope our treatment will decrease trust in (ayurvedic doctors and word of mouth/rumors).

Next, we measure situation-specific trust in sources by presenting respondents with a vignette describing a scenario in which they are seeking emergency advice when a family member falls ill. We then ask respondents, on a scale from strongly trust to strongly distrust, to indicate which sources they would rely on for information in the event of this medical emergency. Our response options include three trustworthy sources (community health center, government-issued health pamphlets or posters, TV interview with a doctor from a leading hospital) and three untrustworthy sources (family stories or remedies, WhatsApp forwards, TV interview with an ayurvedic doctor). The aim of this specific trust measurement was to provide respondents with a concrete situation to consider. Additionally, the distinction between general and situation-specific trust allowed us to separate mediums of news (such as TV or social media) from transmitters, recognizing that our decisions on whom to trust are influenced by a combination of both (transmitter  $\times$  medium).

Lastly, we ask about the factors that are pivotal in fostering trust in a piece of information when encountered specifically on social media. In this section, we delineate several characteristics of messages that respondents could utilize as cues in determining whether to trust them. These included less trustworthy signals such as the number of likes or shares a post has received and whether the sender is from the same community, as well as other signals such as the tone or emotionality of a message.

We find that BIMLI significantly altered respondents’ interactions with and trust in sources overall. There was a notable change in the index ( $SD = 0.21$ ) in the anticipated direction. Ad-

ditionally, we demonstrate that both in generic and specific source discernment, the treatment seemed to operate by diminishing trust in unreliable sources rather than augmenting trust in reliable ones. Lastly, we observed a robust effect of reduced trust in less trustworthy cues when evaluating information online.

#### **4.1.5 Engagement with misinformation**

We measured engagement with misinformation through both attitudinal and behavioral measures. On attitudes, we asked 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 demonstrate the treatment had significant influence on respondents' attitudes towards engagement with misinformation index, although there is notable variation in responses at the item level. First, treated respondents exhibit a marked increase in their self-reported likelihood to abstain from sharing misinformation, even if it originates from close acquaintances. However, they also express a reluctance to correct such misinformation. This finding aligns with prior scholarly observations that cultural facets within the Indian context, such as deference to social norms and respect for interpersonal relationships, may discourage direct confrontation in correcting misinformation (Malhotra and Pearce, 2022). Consequently, it is unsurprising that respondents are hesitant to endorse corrective actions toward friends. Nevertheless, the inclination to refrain from disseminating misinformation indicates a discernible shift in attitudes induced by the treatment. On the question about the importance of fact-checking, we encounter a ceiling effect, with 83% of all respondents affirming its significance, including 82% within the control group. Thus, it is plausible that the treatment had limited capacity to influence beliefs in this regard, given the high baseline endorsement of the importance of fact-checking across both groups.

Finally, we included two behavioral measures. First, we ask respondents to enter a lottery to choose one of two subscriptions to paid information services that we give them for free as

an extra gift for completing the program. Our options include a year-long free subscription to Hindustan, one of India’s oldest and most popular Hindi newspapers providing largely reliable and trustworthy news and information. The second option is a year-long free subscription to Manohar Kahaniyan, a popular monthly Hindi magazine in India, providing short stories on a number of topics including horror, mystery, romance, and suspense. We expect that demand for news and information (over entertainment) will be higher in the treatment group. Second, we tell respondents that we are selecting a group of people to become truth ambassadors and represent their village in the future during times of conflict such as another health epidemic, and that their job would be to help local representatives dispel misinformation that arises during the crisis. We expected that willingness to engage in this effort will be higher for treatment group respondents.

Results demonstrate that we encounter ceiling effects once again. For instance, in the case of the newspaper preference question, nearly all respondents, irrespective of their treatment status, indicated a preference for Hindustan over the alternative option (see Table 4). This trend could potentially signify a widespread preference for this particular news source within the population, as reading newspapers is encouraged by teachers in school and is common among those interested in preparing for competitive exams after school. Alternatively, it might reflect the heightened prominence of Hindustan as a household name.

Table 4: % saying Yes (behavioral measures)

Item	Full Sample	Control	Treatment
Prefer Hindustan paper	82	81	83
Willing to be truth ambassador	90	90	90

## 4.2 Heterogeneous treatment effects

We next discuss heterogeneous treatment effects. We pre-registered an examination of such effects based on pre-test variables identified in the literature as significant in this context. To proxy for motivated reasoning, we estimate interaction effects with partisan identity and attitudes towards non-scientific belief systems (e.g., Ayurveda). To determine if those more connected to the internet respond differently to the treatment, we examine household mobile internet access. We also assess pre-test knowledge of basic science through a battery of questions. Demographically,

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 K, reveal no consistent patterns for heterogeneous treatment effects. Importantly, we do not detect interaction effects for any index with prior mobile internet usage, age, or party identification. These findings are noteworthy since previous research demonstrates that partisanship is both a strong predictor and a factor that can dilute the efficacy of misinformation countermeasures. Our lack of findings in these areas suggests that the mechanisms of belief change in our context may operate through pathways that are not dependent on prior identities or online access to information. Furthermore, they demonstrate that our treatment was robust enough to overcome partisan or identity-based motivated reasoning.

However, we do observe some interaction effects with certain indices for science knowledge, where the treatment appears to help those with higher levels of science knowledge improve their source discernment and engagement attitudes. We also find that prior attitudes towards Ayurveda interact with the treatment to affect engagement attitudes. Nonetheless, these effects do not exhibit a consistent pattern, and they are small with large confidence intervals. Since these effects are observed only with certain interaction variables and specific indices, we cannot rule out the possibility that they are due to chance. Finally, we do not find that respondents in the high spillover versus low spillover strata behaved differently from each other in response to the treatment.

### **4.3 Durability of effects**

Due to the logistical challenges of implementing a large-scale, in-person survey in a low-development context, the duration of our endline survey was 6 weeks long, i.e. data for individual respondents was collected anywhere between 3 days to 6 weeks after the last classroom session. While we were unable to randomly assign the order in which villages were surveyed, we note that since treatment was assigned independent of the villages' background characteristics, we expect that endline survey timing — although dependent on a number of village-level covariates — is independent of treatment status. We use this non-random variation in endline timing to investi-

gate the stability and durability of treatment effects. We first examine the intent-to-treat effects by subgroups categorized by the week of the survey. The analysis spans from week one (the first seven days following the final class for each respondent) through six or more weeks post-intervention (Appendix L16). Findings indicate that several of our results remain stable for up to six weeks after the intervention. Specifically, treatment respondents surveyed in week six are still significantly different from the control group in accuracy and sharing discernment, health preferences, and source discernment. Additionally, the effect sizes observed in later weeks are still larger compared to other similar studies. Attitudes towards engagement with misinformation countermeasures remain significant till week 5, and with behavioral measures, while we do not detect an effect overall, we find that the subgroup interviewed in week 2 was significantly different than control.<sup>14</sup> Next, we look at subgroup ITT effects by plotting the estimated effects in intervals of 3-day bins after the last classroom session (Appendix L17). Collectively, these results demonstrate that the treatment not only produced large and substantial effects but also that these effects did not diminish up to six weeks after the intervention, underscoring its effectiveness in inducing long-term attitudinal changes.

Finally, we also conducted a follow-up survey with a random subset of 2,000 respondents six months after the intervention concluded to assess its long-term persistence and potential network effects. The extended time gap is particularly significant as between both our endlines, India's 2024 general elections took place – a period during which, as shown by previous research, political and partisan attitudes tend to solidify (Michelitch and Utych, 2018). 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, during this second endline survey, we included measures of discernment concerning political stories that gained prominence post-election. We find that treated respondents demonstrated a significantly greater capacity to accurately identify the veracity of these political narratives (0.31 SD). This outcome is especially noteworthy given that our intervention was exclusively focused on health-related content and did not address po-

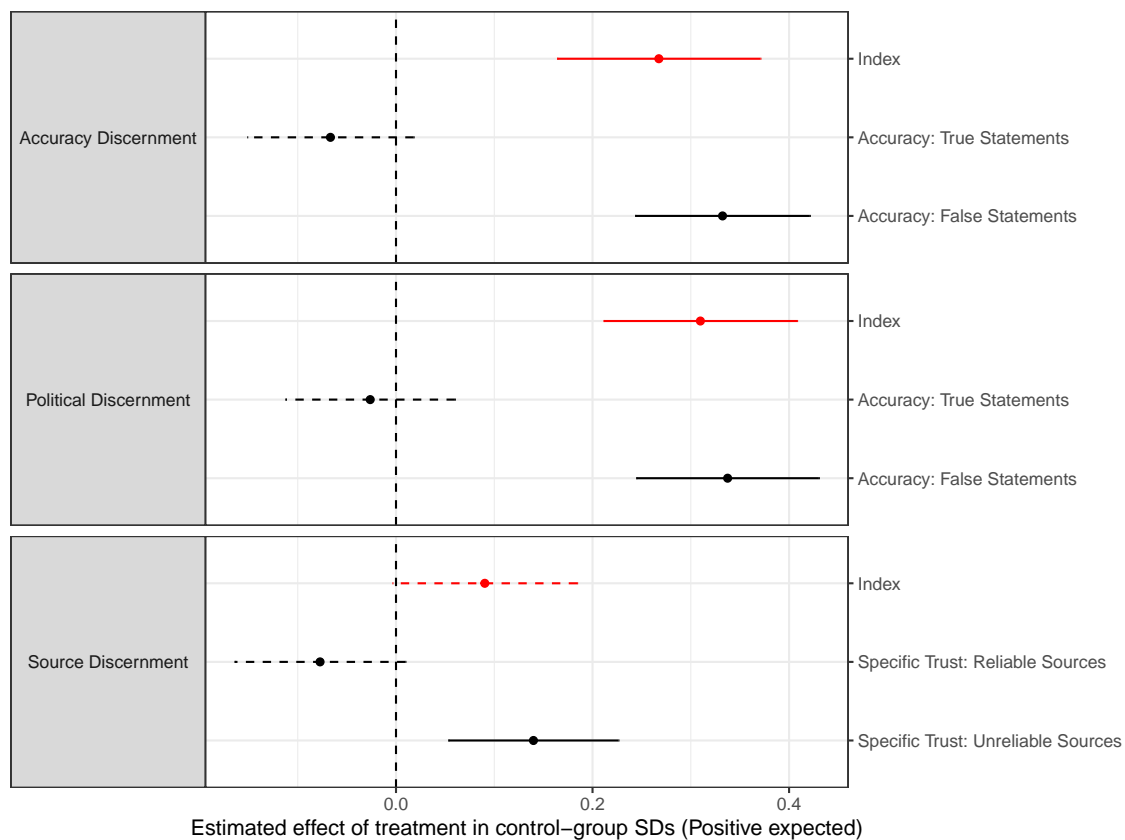
---

<sup>14</sup>We note that the lack of effects in Week 1 i.e. immediately after the intervention is likely because in the initial stages we rolled out the survey slowly and hence our Week 1 samples are substantially smaller than in other weeks.



litical 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: media literacy training, 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. Finally, our second endline also surveyed a random subset of 2000 parents of students who participated in the RCT. Findings demonstrate that parents of treated students are significantly better able to discern true from false information, underscoring the learning mechanism of children affecting their parents' attitudes and behaviors through trickle-up socialization (Dahlgard, 2018).

Figure 7: Effect of assignment to BIMLI treatment on 6-month follow up



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

## 5 Discussion and Conclusion

In this study, we examined the effects of a large-scale, iterative, grassroots media literacy intervention implemented among over 12,000 adolescents in Bihar, India. Collaborating with a local governmental body, we provided media literacy classes to students over the course of half a school year. Intent-to-treat estimates overwhelmingly 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. By the conclusion of the curriculum, participants demonstrated heightened discernment in evaluating information and making decisions regarding the sharing of news items. Notably, the intervention also brought about changes in their health preferences, diminishing reliance on alternative medical approaches to address health concerns. Further, it enhanced their assessment of sources and their ability to discern their credibility. These effects showed continued persistence with a sub-sample of respondents interviewed 6 months after the intervention, as well as spillovers to parents of treated students.

The implications of these findings hold significant importance, especially in light of the prevailing challenges faced by many media 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, where our study took place, typically exhibits low prioritization of education, contributing to a substantial dropout rate in public schools (Muralidharan and Prakash, 2017). Ministry of Education data underscores Bihar's status as one of the poorest performing states in terms of secondary school dropout rates, with 42% of students discontinuing their education before completing the 10th grade. Although we achieved about 70% compliance for each session, the inability to achieve even greater compliance may not solely be attributed to the content of our sessions or the way we ran BIMLI. Rather, it may reflect the inherent realities of sampling from a rural setting characterized by limited state capacity. National Family Health Survey (NFHS) indicators for Bihar consistently underscore its under-performance compared to the rest of the country across various metrics, primarily attributable to extreme state weakness. Within this context of state dysfunction, it is notable that public programs across the board demonstrate weak performance.

Hence, it was not readily foreseeable that a curriculum like BIMLI could yield such signif-

icant outcomes within this context. 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, behavioral intentions, and information consumption patterns. We acknowledge the potential differences in outcomes had the intervention targeted adults, who might have been more susceptible to motivated reasoning or less amenable to change (Badrinathan, 2021; Taber and Lodge, 2006). However, this underscores one of our central arguments: interventions should prioritize populations whose attitudes and behaviors are amenable to modification (Niemi and Jennings, 1991). Young individuals seeking access to quality educational resources in one of India's most impoverished states may represent such a cohort.

Despite the encouraging findings of this study, we acknowledge several limitations. First, the generalizability of our findings is constrained to the scope conditions we identify. Importantly, this includes households that opted to enroll their children in additional classroom hours; it is likely that households that did not self-select into the study initially are substantially different on a number of factors. Moreover, our study achieved a higher degree of compliance with classes through various measures that may depart from a more natural "in-the-wild" roll-out. To address potential biases stemming from variations in state capacity across villages, we undertook several design decisions including engaging external teachers to deliver the BIMLI curriculum, implementing rigorous field monitoring to enhance compliance, and utilizing community libraries rather than schools as classroom venues. These strategies aimed to mitigate the influence of state capacity disparities and enhance the program's reach and effectiveness.

One of the limitations of this work is that such an intervention is costly to set up, and raises the question of whether such an effort-intensive and sustained treatment is necessary to counter misinformation. Indeed, many existing studies use brief fact-checks or primes and find effects on outcomes. To explore this, we examine the number of sessions attended by students and interact compliance with the treatment to assess whether attending more classes – i.e. engaging more intensively with the treatment — alters outcomes. The results, presented in Appendix M, indicate that attending just one session does not significantly impact *any* of the outcomes we measure. This finding aligns with existing research suggesting that one-time media literacy treatments may be insufficient to address the misinformation problem in deeply polarized countries (Badrinathan, 2021). However, we find that for accuracy and sharing discernment, attending two

sessions produces a significant effect compared to the control group, with no additional effect observed when increasing attendance to three or four sessions. In contrast, outcomes related to health preferences, source discernment, and attitudes towards counter-misinformation measures do not shift relative to the control group until after three sessions of attendance. These results suggest two key insights. First, discernment tasks are more easily influenced than other outcomes, yet they require at least two sessions (or three contact hours) to achieve significant change. And in doing so, we detect effect sizes on discernment outcomes that are 2 to 3 times larger than those detected in existing work. Second, outcomes requiring three or more sessions to change are typically attitudes that are more ingrained and pre-date the circulation of some of the misinformation stories measured in our discernment tasks. For instance, the belief that COVID-19 can be cured by cow urine, a story popularized during the pandemic, contrasts with the longstanding belief in non-scientific systems like Ayurveda for curing diseases. Previous research on misinformation indicates that such deep-rooted beliefs are indeed harder to dislodge (Chauchard and Badrinathan, 2024). Moreover, scholarship on polarization and social contact demonstrates that attitudinal changes do not occur easily despite intensive social contact treatments (Scacco and Warren, 2018; Mousa, 2020). Nevertheless, our findings show that deep-rooted attitudes can change following intensive, education-based interventions. These results underscore the need to move away from light-touch interventions and emphasize that altering misinformation and misinformed attitudes requires substantial effort in developing contexts. They also highlight the necessity for sustained interventions focused on learning.

Despite these limitations, the positive effects we find are hopeful, given especially that existing media literacy initiatives in the global south have not been met with success. We attribute this 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\\_docs/PA0215JW.pdf](https://pdf.usaid.gov/pdf_docs/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.
- 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*.
- 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.
- Cao, Jian, Christina M Ramirez, and R Michael Alvarez. 2022. "The politics of vaccine hesitancy in the United States." *Social Science Quarterly* 103 (1): 42–54.

- 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.
- Dasgupta, Aditya, and Devesh Kapur. 2020. "The political economy of bureaucratic overload: Evidence from rural development officials in India." *American Political Science Review* 114 (4): 1316–1334.
- 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.
- Gadjanova, Elena, Gabrielle Lynch, and Ghadaifi 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.
- 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.
- Kahan, Dan M, Ellen Peters, Erica Cantrell Dawson, and Paul Slovic. 2017. "Motivated Numeracy and Enlightened Self-Government." *Behavioural Public Policy* 1 (1): 54–86.
- 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.
- 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.
- Margolis, Michele F. 2018. *From Politics to The Pews: How Partisanship and The Political Environment Shape Religious Identity*. Chicago: The University of Chicago Press.
- 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.



- 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.".
- Persson, Mikael, Karl-Oskar Lindgren, and Sven Oskarsson. 2016. "How does education affect adolescents' political development?" *Economics of Education Review* 53: 182–193.
- Porter, Ethan, and Thomas J Wood. 2019. *False Alarm: The Truth About Political Mistruths in the Trump Era*. Cambridge University Press.
- Rampal, Nikhil. 2023. "Indian states with less than 10 says 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, JNK, and AJ Scott. 1987. "On simple adjustments to chi-square tests with sample survey data." *The annals of statistics*: 385–397.
- 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.
- 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.