

Evaluation Brief

The National Democratic Institute

Global Design, Monitoring, Evaluation & Learning

The Issue

MISINFORMATION & POLARIZATION

NDI has seen the increase in misinformation and polarization create new challenges to every aspect of its democracy assistance, including support for electoral integrity, political party development, democratic governance and inclusive citizen participation.

THE DEMOCRACY CHALLENGE

What kinds of online interventions — separately or in combination — can impact youth capacity to identify misinformation, reduce their uptake and dissemination of misinformation, and mitigate affective polarization across partisan divides?

EVALUATION DESIGN

With funding from the [NED](#) and in partnership with academic researchers from Evidence in Governance And Politics ([EGAP](#)), NDI experimentally tested the impacts of four types of intervention hypotheses; one capacity-based vs. three social identity interventions

Mitigating Online Misinformation and Polarization in Conflict-Sensitive Contexts

Experimental Evidence from Cote d'Ivoire

Misinformation and its spread through social media exacerbates socio-political distance and animosity between citizens (affective polarization). A favorite tool of the aspiring autocrat, misinformation can drive political parties to the extremes, escalate nationalistic rhetoric, divide the electorate and undermine fragile democracies.

In emerging democracies, what motivates the citizenry to consume and assimilate biased media, even when they may know it is false? Is there a psycho-social benefit to assimilating misinformation even when “information consumers” have the capacity to identify false or misleading online information. These were some of the driving questions NDI and its academic partners sought to answer in Côte d'Ivoire . Understanding citizen motivations could illuminate why standard interventions to counter misinformation — e.g., digital literacy, fact-checking, debunking — may fall short in their efficacy. Behind these kinds of ex ante and ex post traditional interventions is the critical assumption that citizens value information accuracy above and beyond the affirmation of their own socio-political identities. But what if this assumption is wrong?

Affective Polarization

The tendency to dislike and distrust members of a political out-group

Capacity vs. Motivated Reasoning: On a practical level, misinformation spreads faster than online fact-checking, making the latter extremely limited. In addition, evaluations of capacity building interventions like digital literacy programs that seek to “inoculate” citizens against the online virality of misinformation have extremely mixed results. In polarized contexts like Côte d'Ivoire individuals identify with certain social categories and may be motivated to sort online information into those social taxonomies, regardless of the veracity of the information itself. That is, in a highly polarized contexts, partisans may be motivated to consume information that affirms their own socio-political identity groups and reject information associated with cross-partisan identities.



The Context

Côte d’Ivoire has long been a polarized society. Two civil wars pitted Southerners against Northerners in a battle to define the “true” Ivorian identity.

The regional cleavages in Côte d’Ivoire have a long history in which relative prosperity attracted migration from neighboring, mostly Muslim countries. From 1922 to 2000 the Muslim populations increased from 6% in the North to 39%. The 2014 census records Christians constituting about 34% of the population compared to 42% Muslim. The country’s first post-independent president, Félix Houphouët-Boigny, maintained peaceful relations within an increasingly diverse population, in large part due to the nation’s long-lasting economic prosperity. But economic crises in the 1980s, followed by Houphouët-Boigny’s death in 1993 brought tensions to the surface.

Two members of Houphouët-Boigny’s government vied to take his place – Henri Bédié who was the former President of the National Assembly, and Alassane Ouattara who was the former Prime Minister. Bédié won and ruled until 1999 when he was overthrown in a coup. The catalyst for the first violent conflict that started shortly thereafter was the passage of a law requiring both parents of a presidential candidate to be born

in Côte d’Ivoire. The law was seen as targeting the exclusion of Alassane Ouattara who was planning to stand in the 2000 election to replace the recently deposited Bédié. Instead, Laurent Gbagbo was elected as the only main opposition candidate able to run in the 2000 presidential elections against the head of the transitional military government.



Ouattara then won the next presidential election in 2010, postponed from 2005 due to the ongoing conflict. While violent conflict in Côte d’Ivoire has mostly subsided, these deep regional divides are now enshrined in electoral politics. Southerners tend to be Christian and supporters of the new opposition party FPI, while Northerners tend to be Muslim, ethnic Mandinke or Barkanabe and supporters of Ouattara’s now-ruling party RHDP.

The Programming

NDI has worked with civil society and political parties in Côte d’Ivoire since 1993, assisting in developing a code of conduct following the 2011 political crises and helping to monitor its compliance during the 2015 and 2016 elections. NDI also played a role in mediating discussions between the government and opposition parties on

a variety of electoral reforms. In 2015 NDI supported the country’s first ever parallel vote tabulation and has continued to assist newly-elected leaders to fulfill their roles. Against a backdrop of conflict and tension, the country has an internet penetration rate of approximately 36%, with more than 23% of its

population using social media, a number that is growing rapidly every year by 8.5%. [A 2020 NDI study](#) showed that Facebook’s most widely shared posts in Côte d’Ivoire were characterized by rampant misinformation that included **assassination rumors, vaccine skepticism, xenophobic hate speech and doxing of political**

opponents. Fake news in Cote d’Ivoire spreads primarily via WhatsApp and Facebook, amplifying political tensions and social divisions particularly around presidential and legislative elections. NDI’s programming focused on social cohesion and mitigating misinformation through digital literacy interventions.

The Experiment

Can social interventions reduce misinformation uptake by reducing the motivated reasoning associated with affective polarization?

The assumption behind digital literacy interventions is that information consumers value accuracy over socio-political identities. However, in a politically contentious context, are information consumers motivated to believe fake news, despite their capacity to identify it?

On the heels of NDI's misinformation study in Côte d'Ivoire, NDI's Global Evaluation & Learning team partnered with academic researchers Dr. Claire Adida, Dr. Jessica Gottlieb and Dr. Richard Moussa to experimentally test four online intervention hypotheses. One hypothesis was capacity-based and three were based on social motivations:

Hypothesis 1: A digital literacy intervention (Capacity) providing information about what misinformation is and how it polarizes society **will not reduce** polarization or the propensity to believe and disseminate misinformation.

Hypothesis 2: An empathy intervention (Motivation) providing individuals with a narrative about the out-group that elicits empathy **will decrease** affective polarization and through it, **reduce** the propensity to believe and disseminate misinformation

Hypothesis 3: A social norms intervention (Motivation) providing individuals with the perception of a diversity of experiences among the in-group toward the out-group **will decrease** affective polarization and through it **reduce** the propensity to believe and disseminate misinformation.

Motivated Reasoning

The phenomenon in cognitive science and social psychology in which emotional biases lead to justifications or decisions based on their desirability rather than an accurate reflection of the evidence.

Hypothesis 4: A popularity intervention that demonstrates positivity (Motivation) **will decrease** affective polarization and through it, **reduce** the propensity to believe and disseminate misinformation.

To test these hypotheses, NDI developed online interventions in partnership with the academic research team, using a two-wave panel survey experiment with Ivorian youth (18-30 years old) in the country's economic capital, Abidjan. Participants were recruited into the study through random-walks in three of Abidjan's 10 communes — Abobo, a majority-

Northerner commune; Prt-Bouet, a majority Southern commune; and Yopopugon, a mixed group commune.

In the first wave of the survey, enumerators administered a questionnaire. Here they randomly assigned participants to listen to one of the audio and/or video treatments. At the end of an initial battery of largely demographic questions, enumerators handed their tablet and headphones to the participant for self-administration of the video and audio intervention(s). Immediately following the intervention, respondents then self-administered a short battery of questions constituting a manipulation check and measurement of intermediate outcomes.

In the second wave, which occurred two to six weeks later, participants completed an online survey through their preferred mode of contact, i.e., Whatsapp, email, or SMS. In the endline survey participants were asked to a) correctly **identify** misinformation; b) their intent to **share** misinformation; and c) answer a battery of questions on feelings of **warmth and trust** toward in-groups and out-groups.

The Findings

Digital Literacy interventions do not change the way individuals consume and share information, while social interventions do, but in unexpected directions.

Correct Identification of Information

- ❖ The **Digital Literacy** intervention had no effect on correctly identifying information
- ❖ The **Empathy** intervention increased the respondent’s ability to correctly identify information
- ❖ The **Norms** intervention decreased the respondents ability to correctly identify information
- ❖ The **Popularity** intervention had no effect on correctly identifying information

Knowingly Sharing Misinformation

- ❖ The **Digital Literacy** intervention had no effect on the respondent knowingly sharing misinformation
- ❖ The **Empathy** intervention had no effect on the respondent knowingly sharing misinformation
- ❖ The **Norms** intervention had no effect on the respondent knowingly sharing misinformation.
- ❖ The **Popularity** intervention decreased the likelihood respondents would knowingly share misinformation

Effect on Polarization

- ❖ On average, neither the capacity treatment nor the motivation treatments had a significant impact on effective polarization. However, the measurement index may have been masking variation within the population.

Aligned and Non-Aligned Groups

Upon further analysis, the data suggest that the motivation interventions may work differently among ethnic groups historically aligned with one side of the North/South cleavage vs. groups that are not aligned, e.g. groups from the central regions.

- ❖ Aligned groups are **more polarized** than unaligned groups at baseline
- ❖ Polarization among **Northerner-aligned groups** is driven by a greater sense of “warmth” toward their own in-group, while **Southerner-aligned groups** are driven by a lower sense of “warmth” toward out-groups
- ❖ The **Norms** intervention decreased out-group warmth among the aligned groups, a potential explanation for the unexpected negative effect on correct identification of information
- ❖ Aligned groups engage in **motivated reasoning** (answer correctly more often when information is aligned with their in-group) while unaligned groups do not.
- ❖ **Norms** interventions increased motivated reasoning, while **Empathy** interventions decreased motivated reasoning, especially among groups aligned with the North/South cleavage.

Average Treatment Effects Across Full Sample

Outcomes of Interest:	Correct Identification of Information	Knowingly Sharing Misinformation	Polarization
Digital Literacy	NULL	NULL	NULL
Empathy, perspective taking	Increase in Ability	NULL	Decrease among least polarized Increase among most polarized
In-group Norms	Decrease in Ability	NULL	Decrease among least polarized Increase among most polarized
Popularity	NULL	Decrease in sharing	Decrease among least polarized Increase among most polarized

Programmatic Takeaways

Social interventions show promise in reducing misinformation uptake by reducing the motivated reasoning associated with polarization

Digital Literacy Has Limited Impacts in the Global South



While there is some evidence that capacity interventions are effective in the Global North, the growing research literature urges skepticism about interventions aimed at correcting information accuracy in non-Western contexts. Positive effects are observed in highly educated samples in the Global South, but not in more representative samples. In contexts where political polarization is driven by social identity attachments, effective interventions need to address the **motivated reasoning** that overrides information accuracy and encourages misinformation uptake in order to

affirm one's group identity.

Empathy and Perspective-Taking Programs Can Reduce Misinformation Uptake



Empathy intervention work on an individual level, encouraging a person to take the perspective of people who don't share their social identity. While the empathy intervention did not mitigate affective polarization, it did **decrease the motivation to believe misinformation**, especially among individuals aligned along the North/South cleavage. This suggests interventions that encourage perspective taking across divides — e.g., popular theatre, radio dramas, written and visual storytelling — may have the potential to reduce uptake and dissemination of misinformation in a

polarized context by encouraging empathy across socio-political divides.

Online Influencers Show Promise in Mitigating Misinformation Uptake



The popularity intervention — listening to an online influencer share positive narratives of altruism and unity — did not decrease polarization along North/South lines. However, it did increase a sense of “in-group warmth” among co-ethnic groups not aligned along the North/South divide, along with **decreasing the likelihood of sharing misinformation**. While the results are not conclusive, they suggest that programs aimed at changing the polarizing narratives of social and political leaders may have a collective impact on reducing the sharing of misinformation.

Attempts to Change In-Group Norms May Have Unintended Negative Impacts



The Norms intervention — listening to members of one's own in-group talk positively about members of an out-group — had the **unexpected, negative impact** of decreasing an individual's capacity to correctly identify misinformation. This suggests that the intervention caused members of the in-group to double-down on their motive reasoning. These unintended, negative consequences underscore the importance of rigorously testing interventions, before going to scale with programs in conflict-sensitive contexts.

Building a Body of Evidence for Democracy Assistance

The National Democratic Institute is a nonprofit, nonpartisan organization working to support and strengthen democratic institutions worldwide through citizen participation, openness and accountability in government. This study was made possible with funding from the the National Endowment for Democracy ([NED](#)). NDI would like to thank Dr. Claire Adida ([UCSD](#)), Dr. Jessica Gottlieb ([U. Of Houston](#)) and Dr. Richard Moussa (ENSSEA) for their leadership and pro bono contributions in designing and implementing this seminal field experiment. In line with the highest ethical standards, this study was [pre-registered](#) (20211117AA) and peer reviewed with the [Evidence in Governance And Politics](#) network, and received IRB approval from the UC San Diego and the University of Houston.



National Democratic
Institute

EvalBriefs

www.ndi.org



Global Design, Monitoring, Evaluation & Learning

NDI believes that evidence for how and why our programs work is key to supporting our democratic partners, capturing promising practice, and holding ourselves accountable for high quality democracy assistance programming overseas. In adherence with the highest professional and ethical standards, NDI believes that the best evaluation methods are those “best-suited” to answer the evaluation questions. Therefore, our methodologies range from randomized controlled trials (RCT) in partnership with leading academic researchers, to participatory evaluations in partnership with local democratic actors.

For more information on this or other evaluative research studies, please contact [Linda Stern, Director of G-DMEL at the National Democratic Institute](#)