LUMS | Chaudhry Nazar Muhammad Department of Economics

A TOOLBOX OF INDIVIDUAL-LEVEL INTERVENTIONS AGAINST Online Misinformation

Policy Brief: October 1st, 2024

The spread of misinformation through media and social networks threatens many aspects of society, including public health and the state of democracies. One approach to mitigating the impact of misinformation focuses on individual-level interventions, equipping social media users and the public with essential tools to discern the authenticity of information, thereby curbing the spread and influence of falsehoods. This policy brief describes review of individual-focused interventions that have been scientifically evaluated worldwide for reducing harm from online misinformation. The findings of this review provide an essential toolbox of interventions covering 9 different types of individual-focused interventions based on a review of 81 scientific papers from across the globe. It provides an accessible, up-to-date scientific knowledge base for policymakers and social media platforms that can inform policy discussions about misinformation countermeasures and platform regulation.

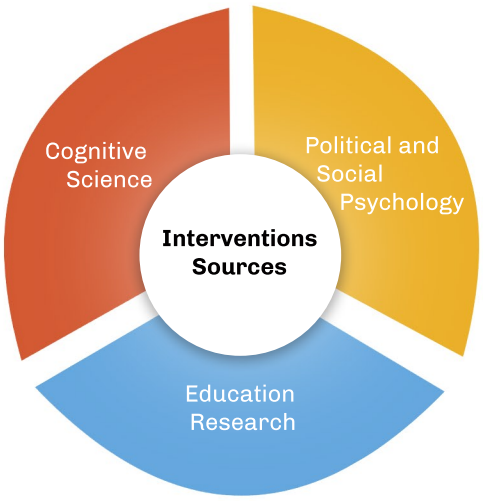
Contributing Author
Ayesha Ali
(along with an international group of experts)

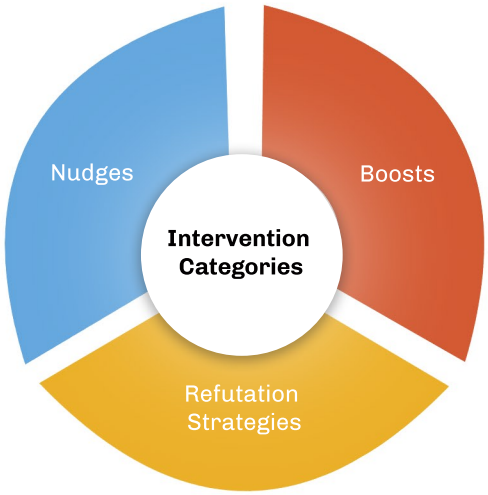
Kozyreva, A., Lorenz-Spreen, P., Herzog, S. M., Ecker, U. K. H., Lewandowsky, S., Hertwig, R., ... Wineburg, S. (2024). Toolbox of individual-level interventions against online information. Nature Human Behaviour, 8(6), 1044–1052. doi:10.1038/s41562-024-01881-0

Introduction

False and misleading online information is a complex and global phenomenon, influenced by both large online platforms and individual and collective behaviour. Misinformation can have far-reaching social, political, and economic impacts by manipulating the attitudes and behaviors of the general public. Online platforms have been struggling to deal with misinformation in multiple ways; highlighting certain content features to end users, restricting content circulation, and by removing false content. Content moderation, however, remains a contentious issue.

Content moderation is a contentious issue and a moral minefield, where the goal of preventing the spread of harmful misinformation can often clash with freedom of expression. A key policy issue, is, thus, how to reduce the spread and influence of misinformation while keeping harsh content-removal measures to a minimum. Recent years have seen a proliferation of interventions aimed at targeting users' competencies and behaviors in a variety of ways.





Different experimental approaches have been employed to test the effectiveness of individual focused interventions to counter online misinformation, ranging from field experiments to online experiments.

This research study represents a joint effort of an international group of experts (30 misinformation researchers from 11 countries and 27 institutions) to collate evidence about the range of misinformation interventions that have been deployed worldwide. The study aims to [1]:

- produce a practical and relevant toolbox of cognitive and behavioral interventions against online misinformation that can be tailored according to the environment and target audience.
- provide a review of main experimental paradigms that have been deployed and can be useful for future research on misinformation interventions.

[1] The research study neither rates interventions nor does it analyze their comparative effectiveness.

The Toolbox

The toolbox organizes interventions into three categories based on cognitive and behavioral mechanisms through which it targets individual reactions and attitudes towards misinformation.

Details of Intervention Categories in the Toolbox

INTERVENTION CATEGORIES	TARGET	DESCRIPTION
Nudging	Behaviour	Uses principles of human psychology to design choice architectures that steer people's decisions, ideally toward a greater individual or public good.
Boosting and Educational Intervention	Competences	Uses human cognition, the environment, or both to help people strengthen existing competences or develop new ones that are useful for coping with a given policy problem.
Refutation	Beliefs	Reduce false beliefs by providing factual information alongside an explanation of why a piece of misinformation is false or misleading
Note: Interventions may fall under more than one category		

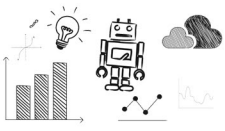
Nine Different Types of Interventions

Nudges

1. Accuracy Prompts
2. Friction
3. Social Norms

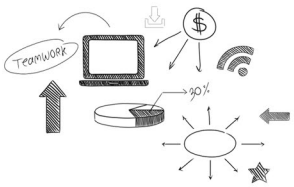
Boosts and Educational Interventions

4. Inoculation,
5. Lateral Reading and Verification Strategies
6. Media-Literacy Tips



Refutation Strategies

7. Debunking and Rebuttals
8. Source-credibility Labels
9. Warning and Fact-checking Labels



For each intervention the toolbox provides:

1. A conceptual overview of interventions. Dynamic table can be accessed **here.**
2. Empirical evidence and the methods used to test interventions can be accessed **here**.

Experimental Paradigms in Testing Misinformation Interventions

Four main experimental paradigms in research on misinformation interventions are also identified:

EXPERIMENTAL PARADIGM	DESCRIPTION	KEY OUTCOME VARIABLES
Misinformation Correction Paradigm	Typically presents people with corrections and measures the impact on their belief in relevant misinformation claims, their claim-related inferential reasoning, and their attitudes towards associated issues. Misinformation correction is one of the earliest paradigms in misinformation research	• Belief and attitude ratings • Reliance on misinformation when responding to inferential-reasoning questions
Headline Discernment Paradigm	Participants evaluate the plausibility, credibility, or veracity of true and false headlines and indicate whether they would be willing to share them. Headline discernment is one of the most widely adopted paradigms in research on misinformation interventions in the last decade.	• Accuracy ratings and people's ability to discern true and false headlines • Measures of sharing intention and sharing discernment.
Technique Adoption Paradigm	This paradigm assesses whether participants successfully learn the skills and strategies required to evaluate information veracity.	• Assessment scores for demonstrating the skills learned during the intervention (e.g., identifying an information source and assessing its credibility)
Field Studies on social media	This paradigm tests lab-based misinformation interventions in the field, often on social media.	• The quality of people's information diets • The quality of what people share on their newsfeeds • People's accuracy or sharing discernment. And can include more variables.

Uses of the Toolbox

The toolbox has the following potential uses in the research and policy landscape:

1. It can provide a starting point for meta-analytic studies, systematic reviews, and studies comparing the effectiveness of different interventions.
2. It can inform efforts to standardize and coordinate methods, thereby increasing the comparability of future results.
3. It highlights important gaps in the available evidence that should be addressed in future studies.
4. It provides accessible and up-to-date scientific knowledge that can inform policy discussions about misinformation countermeasures and platform regulation.
5. The toolbox can also serve as a resource for educational programs, and for individuals who wish to practice self-nudging.
6. The toolbox can help researchers, policymakers, educators, and the public combine interventions to address different aspects of a misinformation problem, especially where a single intervention may have only limited effects.

Limitations and Future Research Direction

01 The interventions in the toolbox act on people's cognition and behavior to reduce either their tendency to share misinformation or the extent to which they are impacted by it. However, there is no simple link between reductions in sharing misinformation at the individual level and the burden and spread of misinformation at the platform level. Individual-level interventions alone are unlikely to fully address the issue of misinformation.

02 User and content-targeted approaches may be limited by how platform design amplifies misinformation. The potential for design changes to reduce misinformation is poorly understood due to a lack of regulated access to platforms' data on their ongoing interventions. Access to this data is necessary for testing types of interventions and design changes at scale.

03 In contrast with individual-level approaches, system-level approaches aim at making the entire online ecosystem less conducive to spreading misinformation—for instance, through platform design, content moderation, high-level regulatory and policy interventions, etc. System-level interventions may be particularly effective and long-lasting. Indeed, they elicit the most confidence among misinformation researchers and practitioners.

The toolbox of interventions against online misinformation, thus, needs to continuously evolve, given the dynamic nature of information environments.

01 It needs to track the extent to which interventions remain relevant and effective, refine existing tools, and develop new ones based on systematic investigations into environmental and individual factors.

02 Methods need to be standardized and coordinated, increasing the comparability of future results.

03 More focus is required to investigate medium- and long-term effects of interventions, explore ways to scale them up (e.g., via school curricula, apps, platform cooperation, and pop-ups), build interventions that reach people across educational backgrounds, and create integrative interventions that empower people to reckon with different types of content.

04 More evidence is needed from Global South and varied cultural contexts.

LUMS ECONOMIC POLICY OUTLOOK SERIES

A Publication prepared by

Rabia Saeed (Research Fellow)

Urwa Fatima (Research Assistant)

Department of Economics, Lahore University of Management Sciences

For information or collaboration,
please reach out at **rabia.saeed@lums.edu.pk**

The source publication can be accessed **here**

Kozyreva, A., Lorenz-Spreen, P., Herzog, S. M., Ecker, U. K. H., Lewandowsky, S., Hertwig, R., ... Wineburg, S. (2024). Toolbox of individual-level interventions against online misinformation. Nature Human Behaviour, 8(6), 1044–1052. **doi:**10.1038/s41562-024-01881-0