

Fundación Maldita.es

**Fact-checking
works: the
evidence on
verification and
the fight against
disinformation**

January 2024

Fact-checking works: the evidence on verification and the fight against disinformation

**Fundación Maldita.es contra la Desinformación:
Periodismo, Educación, Investigación y Datos en Nuevos Formatos**

Avenida del Manzanares 196, Local 2.
28026 - Madrid
SPAIN
CIF: G88519038

Register of Foundations of the Spanish Ministry of Justice: **2330JUS**
EU Transparency Register: **203875243909-78**



According to scientific studies, fact-checking reduces the credibility of misinformation and its dissemination

From time to time it is easy to come across public figures who say that fact-checking “does not work” against disinformation, but in recent years we have dozens of top-level scientific studies that tell us exactly the opposite. The researchers say, for example, that when someone sees a warning that the content they are reading on the Internet is false, that “[effectively reduces the spread of misinformation](#)” and that these warnings “are effective regardless of partisan preferences and other demographic characteristics”.

These results are in line with some data that the digital platforms themselves have made public about their fact-checking programs: according to [Meta](#), when European Facebook users go to share content and see a warning that an independent fact-checker has said that is false, 37% of them decide not to share it. An intervention respectful of freedom of expression, but with a much greater impact than other more invasive ones such as deleting or blocking content (Maldita participates in the Meta verification program).

According to the researchers, fact-checking has proven effective in [correcting false beliefs about COVID-19](#). It has worked even [when a political leader has been denied to his own followers](#), but [politicians also tend to lie less after being corrected by a fact-checker](#). Its effectiveness has been demonstrated [in different countries](#) and situations, while some of the arguments commonly used against fact-checking, such as that it could even be counterproductive, have been [discredited](#) in [different studies](#).

Part of the criticism of the effectiveness of fact-checking is self-serving. Verification requires a lot of human work by properly trained journalists and subject to [international standards of quality and non-partisanship](#). In other words: it is expensive and difficult. It is convenient for many companies to think that they can find artificial intelligence that can tell them if something is true or false, but these systems [fail much more than we can afford](#), and they will never consult human sources: Maldita can call the Marbella town hall to get direct information from a source or go to talk to a witness. The AI cannot.



Fact-checking effectively reduces the spread of misinformation

One of the most common ways citizens encounter fact-checking is through labels on platforms like Facebook or Instagram that warn them that an independent fact-checker says something is false or needs context. At the end of 2023, two researchers from the Massachusetts Institute of Technology (MIT) [reviewed the evidence on its effectiveness](#) obtained in different experiments and concluded that:

- Fact-checking labels reduced people's perception of the truthfulness of a publication by between 13% and 35%
- They reduced the intention to share it between 25% and 46%
- They reduced positive reactions to content, such as “likes”
- More specific and explanatory labels, particularly combined with denials, work even better
- Warnings and labels that come from an independent fact-checker are the most effective

Part of these conclusions are in line with the own evidence of the platforms that use fact-checking warnings and that share them as part of the transparency requirements of the [EU Code of Good Practice against Disinformation](#). Meta says that when European Facebook users go to share a post and are faced with a warning that an independent fact-checker considers it false, [37% decide not to do so](#), as does a slightly higher percentage on Instagram.

Fact-checking is effective: on health, politics and in different countries



Fact-checking has proven effective against very different types of misinformation. In an [experiment with misinformation related to COVID-19 on social networks](#), researchers detected that the accuracy of users who had seen corrections to misinformation had improved by 0.62 out of 5, while that of those who had only seen the misinformation had fallen by 0.13. It is a study designed to replicate the real conditions of a user on the Internet, where different content competes for their attention.

Regarding political fact-checking, whose impact many people believe should be limited by the polarization of society, [four researchers reached the opposite conclusion](#): among the citizens who participated in the experiment, those who had read an article that denied a statement from a political leader had more knowledge of the facts than those who did not, but this improvement was also maintained among those who declared themselves supporters of that politician. Even among his supporters, fact-checking had an effect.

Furthermore, the effect of political fact-checking is not limited to citizens, it also affects politicians themselves. Professors from three Italian universities analyzed whether politicians who had been the subject of denials by fact-checkers lied less afterwards and discovered that this was the case: of the 55 parliamentarians studied, [those who had been “victims” of a denial lied less](#).

Although many studies on the effects of fact-checking focus on the United States and, to a lesser extent, Europe, there are studies that use evidence from other countries in Latin America and Africa. In a [simultaneous experiment in Argentina, Nigeria, South Africa and the United Kingdom](#), similar effects were found in different places, leading the researchers to state that fact-checking “can be a central tool in the fight against misinformation.” Exposure to a denial reduced belief in misinformation by 0.59 out of 5.



The possible adverse effects of fact-checking are very limited, if they exist

Some of the arguments used to attack fact-checking, sometimes even from the academic world, do not have a solid scientific basis. This is the case of “backfire”, the counterproductive effect that could make a debunk increase belief in misinformation. That was the result of an [experiment conducted in 2010](#) that used a politically controversial example (the existence of weapons of mass destruction in Iraq). However, [subsequent researches](#) showed that this counterproductive effect was not repeated. Nonetheless, we usually see how refuted ideas from articles written more than a decade ago continue to cast doubt on whether the work of fact-checkers worsens the effects of misinformation instead of reducing them.

Another common attack is to affirm that the problem of misinformation is being exaggerated. According to this theory, the amount of misinformation is very small and disproportionate concern is being generated around it. However, the [academic articles](#) that are used to justify these types of statements have a basic problem: they make an extremely limited definition of what misinformation is and, therefore, they measure only a small part of the total misinformation to which citizens are exposed.

For example, one of the academic articles that is commonly used to accuse those who fight against disinformation of alarmism is [this one published in Nature](#). Their estimate is that “fake news’ represents only 0.15% of Americans’ daily media diet.” But only what is published by websites labeled as misinforming is considered misinformation, everything else is reliable information. Researchers monitor the impact of the articles on these websites on all types of media but ignore a huge number of formats in which misinformation is spread and, although they are more difficult to narrow down, they are real and affect citizens every day. The fact that we cannot measure all the misinformation that exists beyond that enormously restricted definition they use does not mean that it is not real.

The so-called “implicit truth effect” is also often considered proven, which says that when some social media content is labeled as false, the public interprets that all the rest of the content is true. However, there are studies like [this one published in the journal Political Behavior](#) that reach very different conclusions and state that “exposure to the labels “Disputed” or “Rated as false” did not affect the perceived accuracy of headlines not labeled as true or false”.

What fact-checkers really do:



Sometimes, many of those who argue that fact-checking “does not work” against disinformation do not fully understand the work of fact-checkers today. As [this article by Peter Cunliffe-Jones and Lucas Graves](#) said, in addition to monitoring disinformation and producing denials in different formats, fact-checking organizations like Maldita do many more things: educational campaigns, early detection of disinformation, technological and AI tools, public policies, community creation or collaboration with academics.

In addition to all this, their monitoring and denial work continues to be not only relevant but also the essential basis for many other activities against disinformation: artificial intelligence systems that identify the language of disinformers are trained on databases that the fact-checkers have certified as false content, and [pre-bunking campaigns](#) or the most successful media literacy initiatives are prepared based on the debunks made by the fact-checkers.

Saying that fact-checking “does not work” against disinformation, as is still heard from time to time, is simply inaccurate. Some of those who defend it simply do not have adequate information or maintain an idea of what fact-checkers do that is anchored a decade ago. Others, of course, are interested in maintaining that narrative for business reasons.

At Maldita, we believe that fact-checking will not end disinformation on its own because 1) nothing can completely solve the problem of disinformation, there is no magic method, disinformation has always existed, but we can help people; and 2) because many different and effective interventions are needed to fight it. Those who accuse small fact-checking organizations (many non-profit, like Maldita) of not having enough impact should, firstly, know our work better and secondly, recognize our contributions and their evident impact on society.