



Review

Misinformation warning labels are widely effective: A review of warning effects and their moderating features

Cameron Martel¹  , David G. Rand^{1 2}

[Show more](#) 

 Share  Cite

<https://doi.org/10.1016/j.copsyc.2023.101710> 

[Get rights and content](#) 

Abstract

There is growing concern over the spread of misinformation online. One widely adopted intervention by platforms for addressing falsehoods is applying “warning labels” to posts deemed inaccurate by fact-checkers. Despite a rich literature on correcting misinformation *after* exposure, much less work has examined the effectiveness of warning labels presented concurrent with exposure. **Promisingly, existing research suggests that warning labels effectively reduce belief and spread of misinformation.** The size of these beneficial effects depends on how the labels are implemented and the characteristics of the content being labeled. Despite some individual differences, recent evidence indicates that warning labels are generally effective across party lines and other demographic characteristics. We discuss potential implications and limitations of labeling policies for addressing online misinformation.

Examples of Misinformation Warning Labels

