Mastering critical thinking skills is strongly associated with the ability to recognize fakeness and misinformation

Simone Redaelli [1], Nikola Biller-Andorno [1],
Sophie Gloeckler [1], Jessica Brown[2], Giovanni Spitale [1,3]*
and Federico Germani [1]*

[1]Institute of Biomedical Ethics and History of Medicine, University of Zurich, Zürich, Switzerland,
[2]Trinity College Dublin, Dublin, Ireland, [3]Centre for Medical Ethics, University of Oslo, Oslo, Norway

Introduction: In the current era of digital information overload, individuals are inundated with content of varying quality and truthfulness. Critical evaluation of such content is essential to distinguish between legitimate information and misinformation or fakeness. Despite this need, there is limited empirical evidence on the effectiveness of training in critical thinking skills for enhancing such discernment.

Methods: This study operationalized critical thinking into six measurable concepts: causation and correlation, independent data and replicates, reproducibility, credibility of sources, experimental control, and statistical significance. A pre-registered randomized controlled trial was conducted involving educational video interventions aimed at improving participants' understanding and application of these concepts. Participants were evaluated based on their ability to identify fake tweets and misinformation both before and after the intervention.

Results: A strong correlation was found in the pre-intervention phase between mastery of critical thinking concepts and the ability to identify misinformation and fakeness. However, the video-based intervention did not significantly enhance critical thinking skills nor improve the participants' accuracy in identifying misinformation compared to the control group. The intervention's inefficacy was consistent across various demographic and educational backgrounds.

Discussion: The findings suggest that while mastery of critical thinking is associated with greater resilience against misinformation, current educational interventions—such as short video lectures—are insufficient. There is a pressing need to develop and empirically validate more effective, possibly interactive, training modalities that can foster misinformation-specific critical thinking skills in the general population.

KEYWORDS

critical thinking, fakeness, misinformation, scientific literacy, infodemics, disinformation, education

1 Introduction

There is an unparalleled amount of information accessible on the internet, including web-based encyclopedias, institutional websites, media outlets, and social media platforms (Zarocostas, 2020). This accessibility presents numerous learning opportunities, granting individuals access to a diverse range of knowledge. However, many online users face the challenge of assessing the truthfulness of the information they encounter (Zimmerman et al., 2022). The World Health Organization (WHO) coined the term "infodemic"; although originally referring to phenomena occurring during public health emergencies such as pandemics, the "infodemic" is a construct that can effectively describe the overabundance of true and false information, and its rapid circulation, during events not only directly related to public health, but also garnering widespread attention, such as political elections (Vosoughi et al., 2018; World Health Organization, 2020). The increased spread of and facilitated access to health-related misinformation on social media brings multiple adverse outcomes for public health, including misinterpretation of scientific evidence, polarization of opinions, heightened levels of fear and panic and, ultimately, reduced accessibility of valuable health services (Borges Do Nascimento et al., 2022). To respond effectively to misinformation, it is therefore necessary to undertake multisectoral interventions, encompassing the foundation of digital and scientific literacy among the general public (Borges Do Nascimento et al., 2022; Howell and Brossard, 2021). According to the OECD PISA 2025 science framework, a person should be capable of mastering three domain-specific competences in order to comprehend, participate in thoughtful discussions, make informed choices, and take action on science-related issues: (a) explain phenomena scientifically; (b) construct and evaluate designs for scientific enquiry and interpret scientific data and evidence critically; (c) research, evaluate and use scientific information for decision making and action (OECD, 2023). Such skills, in turn, require scientific knowledge (OECD, 2023). This definition is, at least in part, in agreement with past characterizations that have generally defined scientific literacy as the capacity to comprehend the natural world, grasp the interconnection between science and society, and solve concrete issues through the application of scientific concepts (Cavagnetto, 2010; National Academies of Sciences et al., 2016; Rubba and Andersen, 1978; Valladares, 2021). However, for scientific competence to be a realistic goal for the general public – who must navigate infodemics and differentiate accurate information from misinformation often without specialized scientific knowledge – it should be supported by the development of linguistic and cognitive abilities. These include metacognition (self-awareness of one's own thinking), as well as critical reasoning skills (Cavagnetto, 2010). Indeed, the United States National Academies of Sciences, Engineering, and Medicine underscore that scientific literacy also entails applying foundational literacy skills within a specific domain (National Academies of Sciences et al., 2016). Foundational literacy commonly includes the ability to write words and sentences, processing words and language in oral contexts, use academic vocabulary and language structures, and have the knowledge base required for comprehension of non-technical texts about multiple topics, including science (National Academies of Sciences et al., 2016). Some authors have argued that the ability to extrapolate

meaning from text is a key feature of scientific literacy skills (National Academies of Sciences et al., 2016; Norris and Phillips, 2003; Shaffer et al., 2019). Furthermore, recent work demonstrated that highlighting typical features of misinformation which are of semantic and grammatical nature helps users recognize misinformation (Li et al., 2022), while the detection of logical fallacies and misleading rhetorical techniques serve as red flags of misinformation across various domains, including COVID-19 (Musi and Reed, 2022), climate change (Zanartu et al., 2024), and vaccination (Schmid and Betsch, 2019). Research in the field of inoculation theory has demonstrated that early exposure to logical fallacies can enhance one's ability to detect faulty reasoning and arguments, enabling individuals to later identify them as misleading within misinformation (Biddlestone et al., 2023; Cook et al., 2017). Although this approach enhances misinformation detection through logical and critical reasoning, its effectiveness is limited to content or strategies introduced preemptively. This strategy does not equip the public with a broad-spectrum tool that accommodates for future changes in the misinformation landscape, rather, arguably inoculation serves as a containment strategy, that needs to continuously adapt to counter evolving misinformation tactics.

Despite being recognized by both the UN (Organización de Naciones Unidas, 2018) and UNESCO (Sabzalieva et al., 2021) as essential for achieving the Sustainable Development Goals – and therefore a priority for educational institutions worldwide – critical thinking remains a complex and difficult-to-define concept (Andreucci-Annunziata et al., 2023; Yu and Zin, 2023). There is no single definition that encompasses its multifaceted nature, and there is little agreement on the most effective ways to teach, nurture, or assess it in educational settings (Andreucci-Annunziata et al., 2023; Bates et al., 2025; Prokop-Dorner et al., 2024; Tiruneh et al., 2014; Yu and Zin, 2023). Problematically, some evidence suggests that students do not become significantly more skilled as critical thinkers over the course of their education (Andreucci-Annunziata et al., 2023; Bates et al., 2025; Tiruneh et al., 2014), raising questions about our ability to prepare future citizens to critically evaluate an increasingly complex information ecosystem (Tiruneh et al., 2014) and the efficacy of current educational approaches (Bates et al., 2025; Bhuttah et al., 2024). Indeed, systematic reviews reveal a lack of consensus on effective teaching strategies (Andreucci-Annunziata et al., 2023; Tiruneh et al., 2014; Zeng and Ravindran, 2025), with a growing preference for problem-based learning (Prokop-Dorner et al., 2024; Yu and Zin, 2023), which emphasizes activities on critical thinking development, the integration of digital technologies, and the use of tools oriented toward critical thinking improvement (such as guiding questions or concept mapping), all while incorporating discipline-specific knowledge (Yu and Zin, 2023). On this basis, building on the initial proposition by Fitzgerald (1997), we decided to rethink the concept of critical reading and reasoning as being underpinned by a set of operationalizable, teachable, and trainable critical thinking skills that could empower the public to autonomously navigate the information landscape. Although systematic evaluations of the skills targeted in critical thinking training remain poorly explored (Willingham, 2020), we leveraged a few qualitative studies available in the literature to investigate how operationalized critical thinking could contribute to building scientific literacy and resilience against misinformation: one study described critical thinking as the ability to judge the accuracy of

data, the control of variables, the credibility of sources, and the validity of inferences (Vieira et al., 2011), while another study defined critical thinking skills as the ability to recognize "scientific evidence", evaluate the credibility of sources, find strengths and weaknesses in research design (such as concepts of bias, sample size, randomization and experimental control), acknowledge the role of statistics in supporting data uncertainty, and understand graphical representations of data (Gormally et al., 2012). Intriguingly, one study found that debunking misinformation containing invalid inferences – such as drawing causal conclusions from merely correlational evidence – can effectively strengthen resilience against misleading representations of scientific findings in the media (Irving et al., 2022). Additional research suggests that teaching individuals to evaluate the credibility of sources, such as assessing the quality of URLs or the presence of supporting evidence for claims, can enhance resilience against fake news on social media (Soetekouw and Angelopoulos, 2024). Based on the presented definitions and applications, we argue that critical thinking skills can be understood as the conceptual foundation of validated practices routinely employed by researchers using the scientific method. If it is possible to extract operationalizable concepts from these practices, it is reasonable to assume that the general public could learn scientific and critical reasoning in the same way that scientists apply it, without the need to be knowledgeable about any specific scientific topic and knowledge domain (e.g., medicine, biology, etc.) (Redaelli, 2020).

With this scope in mind, we operationalized critical thinking skills into the following set of measurable concepts potentially useful for reflecting the above practices, developing scientific literacy and robust defenses against misinformation: (1) causation and correlation; (2) independent data and replicates; (3) reproducibility; (4) credibility of sources; (5) experimental control; (6) statistical significance (for a detailed description, see section "2.1 Definitions").

The aim of this large, pre-registered, randomized study is to evaluate whether: (1) participants' pre-intervention understanding and correct application of the aforementioned 6 critical thinking skills concepts correlates with an ability to identify fakeness (false tweets), here referred to as fabricated, accurate or inaccurate information that mimics media content and graphical appearance (Lazer et al., 2018) – and misinformation; and (2) participants who receive critical thinking training through educational videos, compared to those who do not, demonstrate improved understanding and accurate application of the 6 critical thinking skills concepts, as well as enhanced ability to identify misinformation and fake contents online.

2 Materials and methods

We registered the protocol of this study before initiating the data collection. The pre-registration can be found on OSF: https://doi.org/10.17605/OSF.IO/Y2674.

2.1 Definitions

For the scope of this study, we referred to "critical thinking" as the ability to understand and master the following 6 critical thinking skills concepts:

- Causation and correlation: causation describes a situation by which one action, process or condition – the cause – participates in the generation of a second action, process or condition – the effect. The cause determines, at least partially, the effect, and the effect is, at least partially, determined by the cause [see Supplementary File 1 in the study's OSF repository (Spitale et al., 2022) for practical examples]. Correlation indicates that two variables are associated or move together in some way, but it does not imply that one causes the other [see Supplementary File 1 in the study's OSF repository (Spitale et al., 2022) for practical examples].
- Independent data and replicates: a conclusion is reliable and reproducible if it is drawn on the basis of empirical data collected from independent and replicated observations. Independent observations are different observations made by independent observers that lead to similar conclusions, strengthening the reliability of findings. Replicates are repeated measurements or experimental trials conducted under the same conditions to ensure reliability and reproducibility of results [see Supplementary File 2 in the study's OSF repository (Spitale et al., 2022) for practical examples].
- Reproducibility: it refers to the ability to obtain consistent results when, e.g., an experiment, study, or analysis is repeated by e.g., independent researchers, using the same methods, data, and conditions [see Supplementary File 3 in the study's OSF repository (Spitale et al., 2022) for practical examples].
- Credibility of sources: it refers to the trustworthiness, reliability, and authority of the information provided by a source. It is determined by factors such as the source's expertise, accuracy, objectivity, and consistency. A credible source is one that is well-researched, supported by evidence, free from bias or conflicts of interest, and widely recognized as reputable by experts in the field [see Supplementary File 4 in the study's OSF repository (Spitale et al., 2022) for practical examples].
- Experimental control: it refers to the systematic management of variables in an experiment to ensure that only the factor being tested is influencing the outcome. This helps eliminate confounding variables and biases, making the results more reliable and allowing for valid conclusions about cause-and-effect relationships [see Supplementary File 5 in the study's OSF repository (Spitale et al., 2022) for practical examples].
- Statistical significance: it refers to the likelihood that a result or relationship observed in a study is not due to random chance but rather reflects a true effect [see Supplementary File 6 in the study's OSF repository (Spitale et al., 2022) for practical examples].

Throughout the manuscript, we consistently employed the terminology of "real" versus "fake" tweets and "accurate information" versus "inaccurate information" (i.e., misinformation). Real tweets are those generated by existing Twitter users, while fake tweets are those artificially generated with TweetGen (see section "2.4 Generation of tweets"). Both real and fake tweets can contain either accurate or inaccurate information (i.e., misinformation). As for the definition of fakeness, we primarily refer to Lazer et al. (2018), in which fake news is defined as "fabricated information that mimics news media content in form

but not in organizational process or intent" (Lazer et al., 2018). However, we adopt a modified version of this definition. Here we define "fakeness" more broadly, as any fabricated information that mimics media content (including news media content) in form, regardless of its accuracy. This means that an entirely accurate piece of information can still be presented within a deceptive or misleading framework – such as being posted on a fraudulent BBC social media profile. Conversely, false information can also be disseminated through the same fraudulent platform. Both cases fall under our definition of fakeness: the first instance represents fakeness combined with accurate information, while the second represents fakeness combined with misinformation. Regarding the definitions of accurate information and misinformation, we ground ourselves in the current scientific knowledge of and consensus on the topics and information under scrutiny, considering as "accurate" the information based on scientific consensus. To steer clear of dubious and debatable cases, which may be prone to personal opinions and interpretations, we exclusively analyzed and incorporated into our survey those tweets containing information that can be clearly and unambiguously categorized as true (correct information) or false (incorrect information). Notably, if a tweet contained partially incorrect information, indicating the presence of more than one piece of information with at least one being incorrect, then it was classified as false. We acknowledge the diverse definitions of disinformation and misinformation, but we adhere to an inclusive definition, which encompasses false information (including partially false information) and/or misleading content (Roozenbeek et al., 2023).

2.2 Design of the study

Our study was designed as a randomized trial. Before randomization, all participants were tested for their skills in: A) Mastering of critical thinking skills concepts; B) Misinformation recognition (inaccurate vs accurate information); C) Fakeness recognition (fake vs real tweets) – meaning the identification of fabricated, accurate or inaccurate information, that mimics media content and graphical appearance (Figure 1A). To test A), all participants were presented with multiple choice questions designed to assess their ability to mastering the following 6 critical thinking skills concepts: causation and correlation; independent data and replicates; reproducibility; credibility of sources; experimental control; statistical significance. Each participant answered 3 randomized questions (from a pool of 6) related to each concept. In total, participants responded to 18 multiple choice questions (total score 0–18) (Figure 1A), randomized from a pool of 36, with answers in a randomized order.

To test B) and C), all participants were shown 10 tweets (randomly selected from a pool of 20). For all 10 tweets, they assessed whether they were fake or real (total score 0–10) and whether they were accurate or contained misinformation (total score 0–10) (Figure 1A).

In the second phase of the trial, participants were randomly assigned 1:2 to two arms: a non-intervention group and an intervention group (for the number of patients and percent distribution per group, see Supplementary Table 1). The intervention group was exposed to a set of mini video lectures (Supplementary Files 1–6) intended to explain the meaning and

correct use of the 6 critical thinking skills concepts tested in A) (i.e., intervention). The passive, non-interactive videos ranged from 3 to 7 min in length and followed a consistent structure: a brief infographic-based theoretical introduction with a voiceover explaining the concept being addressed; illustrative tweet examples accompanied by a voiceover commentary; and concluding take-home messages summarizing the key points to retain. Participants could complete the intervention in a single session, although they had the option to pause the videos and complete the intervention later. To limit the intervention duration, participants only watched videos related to specific critical thinking concepts if their pre-intervention score for that concept was below 3. For instance, if a participant scored 1 or 2 on "Causation and correlation", they viewed the corresponding video. If they made no mistakes (score = 3), no video was shown for that concept. Participants who made no mistakes in the pre-intervention phase and scored 18/18 on critical thinking concepts were shown a random video if assigned to the intervention group. The 1:2 randomization was chosen due to the anticipated low survey completion rate in the intervention group, attributed to the survey's length resulting from the video-based intervention.

After the intervention, participants in both arms were re-tested for their skills in A), B) and C): they were presented with the 18 questions, as well as the 10 tweets, from the dataset that they had not evaluated during the pre-intervention phase (Figure 1B). Both in the pre-intervention phase and in the post-intervention phase, participants received:

- A score of 0–10, based on their ability to recognize whether the 10 tweets were fake or real (fakeness score).
- A score of 0–10, based on their ability to recognize whether the 10 tweets were accurate or contained misinformation (misinformation score).
- A score of 0–18, based on their understanding of the concepts of causation and correlation, independent data and replicates, reproducibility, credibility of sources, experimental control, statistical significance (critical thinking skills score).

The duration of the survey ranged from 20 min to approximately 1 h, depending on the participant's speed, on whether they were assigned to the non-intervention group or the intervention group, and on how many videos they had to watch. Participants were unable to skip the videos until at least half of their duration had elapsed. Participants were restricted to one attempt at the survey using cookies and IP tracking: those who had already fully or partially submitted the survey were prohibited from accessing it again. Further details are provided in the section "2. Methods" of the study's OSF repository (Spitale et al., 2022).

2.3 Programming of the survey

We programmed a survey with Qualtrics to collect demographics (Supplementary Tables 1–6), display the tweets, multiple choice questions, and video lectures to participants, and collect their assessments. Participants were asked for consent to participate in the study. The Qualtrics survey structure is available in the study's OSF repository (Spitale et al., 2022).
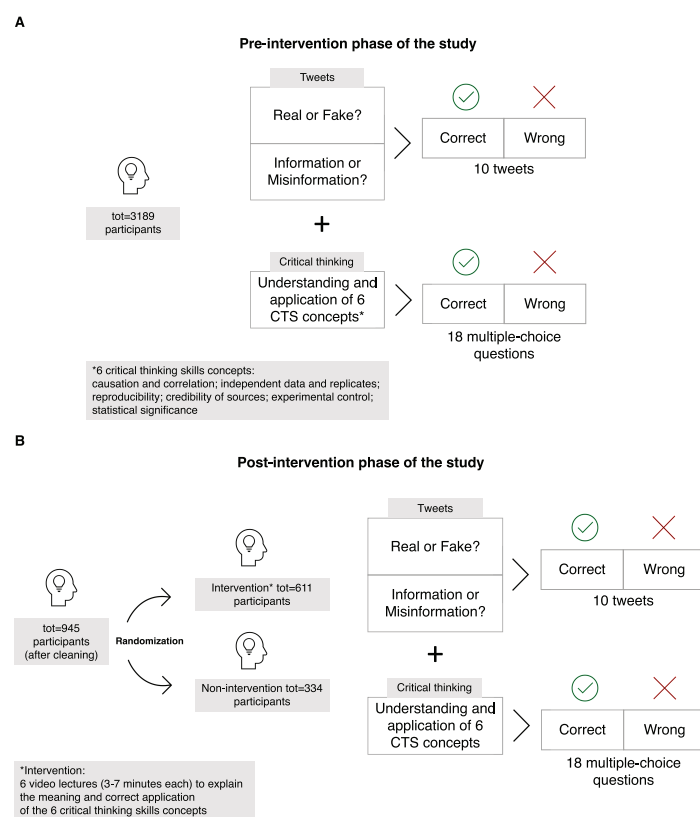
FIGURE 1
Study design. We designed a randomized study to assess whether mastering critical thinking skills correlates with the ability to recognize fakeness (fake tweets) and misinformation. **(A)** Study design of the pre-intervention phase. All study participants (n = 3189) completed a survey where they were asked to assess whether 10 tweets were fake or real and whether they contained accurate or inaccurate information and answer 18 multiple-choice questions that evaluated their understanding and correct application of the 6 critical thinking skills concepts under scrutiny. **(B)** Study design of the post-intervention phase. After cleaning and upon randomization, participants were divided in an intervention group (n = 611) and in a non-intervention group (n = 334). The intervention group underwent an intervention, consisting in 6 video lectures (3–7 min each), aimed at training participants' understanding and application of the 6 critical thinking skills concepts under scrutiny. To limit the intervention duration, participants only watched videos related to specific critical thinking concepts if their score for that concept was below 3/3. Participants who made no mistakes and scored 18/18 on critical thinking concepts were shown a random video if assigned to the intervention group. After the intervention, participants in both groups completed a new survey where they were asked to assess whether 10 new tweets were fake or real and whether they contained accurate or inaccurate information, and answer 18 new multiple-choice questions that evaluated their understanding and correct application of the 6 critical thinking skills concepts under scrutiny. CTS, critical thinking skills.

2.4 Generation of tweets

Tweets were chosen to be a mix of fake and real tweets (see section "2.1 Definitions"), containing accurate information or misinformation (pool of 20 tweets, of which 10 fake and 10 real). More information about the tweets is available for scrutiny and replication in the study's OSF repository and in Supplementary File 7 (Spitale et al., 2022).

2.5 Test survey with convenience sample and pilot testing

We initially conducted a test survey to evaluate and assess the study's potential and to determine whether the ability to identify misinformation and fakeness correlated with mastering critical thinking skills concepts. The survey structure was the same as that of the survey implemented during the pre-pilot and pilot phases of the study. Details are provided in the "Test survey with

convenient sample" section of the study's OSF repository (Spitale et al., 2022). We later conducted a pre-pilot study to verify the technical integrity of the survey's programming and interface. As no issues were identified, we combined the data collected during the pre-pilot study with the data obtained during the pilot study. The pilot study was distributed via a Facebook ads campaign. Details are provided in the "Pilot study" section of the study's OSF repository (Spitale et al., 2022).

2.6 Data collection

We distributed the survey via 3 independent platforms: Facebook ads, mTurk, and the mailing list of the University of Zurich (UZH). All the data, in anonymized form, are available in the study's OSF repository (Spitale et al., 2022). Due to the nature and design of this study it was not possible to conduct a power analysis. Our considerations on the definition of the study's sample size based on the test with a convenience sample are fully detailed in the study's OSF repository (Spitale et al., 2022).

2.7 Analysis

Scoring and analysis were conducted in Python and structured in a Jupyter notebook. The code took the results of the Qualtrics survey as input and generated the files needed for the analysis as output. The code is available for scrutiny and replication in the study's OSF repository (Spitale et al., 2022).

2.8 Cleaning

We included an attention test in the survey to filter results based on participants' attentiveness. For Facebook and email users who failed the attention test, participants were allowed to continue the survey, although their data were not included in the analysis. This approach aimed to encourage them to share the survey link further (with their friends, family, and personal contacts). For mTurk participants, the attention test determined whether participants could continue the survey or not. This strategy helped increase the cost-efficiency of the campaign, ensuring that mTurk Workers who did not pass the test could not receive the code and were not paid. The attention test asked participants whether they identified as vegetarians, explicitly instructing them to skip the question (not to respond) and proceed with the survey (Spitale et al., 2022). It was positioned after the demographics section and before the critical thinking and misinformation tests. The survey structure is available in the study's OSF repository (Spitale et al., 2022). mTurk workers received equal compensation regardless of the time taken to complete the survey; however, those who completed it too quickly (speeders) were not compensated.

2.9 Inferential statistics

For correlation analyses, we performed a Pearson's test. To compare differences between independent groups not normally distributed, we performed a Mann-Whitney U test. To test the effect of independent variables on dependent variables, we performed a MANOVA.

3 Results

3.1 Design and demographics of the study

We designed a randomized trial to investigate whether mastering critical thinking skills concepts correlate with the ability to recognize fakeness (fake versus real tweets) and misinformation (inaccurate versus accurate information) (Figure 1). Overall, we recruited 3,189 participants (mTurk = 436; Facebook = 981; UZH mailing list = 1772). We excluded responses of participants who did not complete the survey, did not give the consent, did not pass the attention test or took shorter than 5 min to complete the survey. Based on these criteria, a total of 2,244 participants (mTurk = 221; Facebook = 806; UZH mailing list = 1217) were excluded; therefore, 945 participants (mTurk = 215; Facebook = 175; UZH

mailing list = 555) were included in our analysis (Supplementary Table 6). The high number of exclusions was not solely due to the survey length, but primarily resulted from our strict adherence to the criteria used for quality control. Pilot testing confirmed that a minimum of 5 min was necessary for participants to provide thoughtful responses, and the attention test further ensured that only engaged participants were included. Although this approach led to a considerable number of exclusions, it was essential for maintaining the integrity and reliability of our findings. Upon randomization 1:2, 334 participants were randomly assigned to the non-intervention group, while 611 participants were assigned to the intervention group (Figure 1B and Supplementary Table 1). Most of the participants were from Switzerland, the United States of America, the United Kingdom, Germany and Italy (Supplementary Table 2), with more males than females (Supplementary Table 3); there was a high representation of people between 18 and 25 years old and between 26 and 40 years old (Supplementary Table 4) and of participants holding a bachelor's degree (or equivalent) (Supplementary Table 5).

3.2 Mastering critical thinking skills concepts strongly correlates with the ability to recognize fakeness and misinformation

In the pre-intervention phase of the trial, mastering critical thinking skills strongly correlated with the ability to recognize fakeness ($R = 0.20$, $p < 0.0001$), misinformation ($R = 0.43$, $p < 0.0001$), or the combination of fakeness and misinformation ($R = 0.40$, $p < 0.0001$) (Table 1A). The degree of correlation between mastering individual critical thinking skills concepts and the ability to recognize fakeness, misinformation or the combination of fakeness and misinformation varied (Table 1B). Causation and correlation emerged as the most relevant operationalized concepts for identifying fakeness and misinformation ($R = 0.35$, $p < 0.0001$), followed by independent data and replicates ($R = 0.29$, $p < 0.0001$), experimental control ($R = 0.29$, $p < 0.0001$), reproducibility ($R = 0.27$, $p < 0.0001$), statistical significance ($R = 0.22$, $p < 0.0001$) and credibility of sources ($R = 0.22$, $p < 0.0001$) (Table 1B). These results suggest that understanding critical thinking skills – as operationalized in this study – is an indicator of the ability to recognize fakeness and misinformation.

3.3 The intervention design is ineffective in training critical thinking skills concepts and does not improve our study participants' ability to recognize fakeness and misinformation

In the post-intervention phase of the trial, we investigated whether our intervention design was effective in training critical thinking skills concepts (for details, see "2. Materials and methods").

We calculated the difference in mean total critical thinking skills scores before and after the intervention for both the

TABLE 1 Mastering critical thinking skills concepts strongly correlates with the ability to recognize fakeness, misinformation, and the combination of fakeness and misinformation.

	Fakeness		Misinformation		Fakeness + Misinformation	
	Pearson's R	*p*-value	Pearson's R	*p*-value	Pearson's R	*p*-value
(A) Pearson's correlation coefficients (R) correlating total critical thinking skills score with fakeness score, misinformation score, and the combination of fakeness score and misinformation score.						
Total critical thinking skills	0.20	<0.0001 (CI: 0.14;0.26)	0.43	<0.0001 (CI: 0.38;0.48)	0.40	<0.0001 (CI: 0.35;0.45)
	Fakeness		Misinformation		Fakeness + Misinformation	
	Pearson's R	*p*-value (CI)	Pearson's R	*p*-value (CI)	Pearson's R	*p*-value (CI)
(B) Pearson's correlation coefficients (R) correlating individual critical thinking skills scores with fakeness score, misinformation score, and the combination of fakeness score and misinformation score						
Causation and correlation	0.20	<0.0001 (0.14;0.26)	0.34	<0.0001 (0.28;0.40)	0.35	<0.0001 (0.29;0.40)
Independent data and replicates	0.14	<0.0001 (0.07;0.20)	0.32	<0.0001 (0.26;0.37)	0.29	<0.0001 (0.23;0.35)
Experimental control	0.15	<0.0001 (0.09;0.21)	0.30	<0.0001 (0.24;0.36)	0.29	<0.0001 (0.23;0.34)
Reproducibility	0.12	<0.0001 (0.05;0.18)	0.30	<0.0001 (0.24;0.36)	0.27	<0.0001 (0.21;0.32)
Statistical significance	0.14	<0.0001 (0.08;0.21)	0.20	<0.0001 (0.14;0.26)	0.22	<0.0001 (0.16;0.28)
Credibility of sources	0.06	=0.06 (0.00;0.12)	0.28	<0.0001 (0.22;0.34)	0.22	<0.0001 (0.16;0.28)

p-value > 0.05, not significant. CI, confidence interval.

intervention and non-intervention groups. When the non-intervention was compared with the intervention group, we found no meaningful nor statistically significant difference between scores (Non-intervention 0.005 versus Intervention 0.018, p = 0.06, Table 2). Furthermore, we tested whether our intervention design was effective in training any individual critical thinking skills concepts. However, we found no meaningful nor statistically significant difference between the intervention and non-intervention groups for any of the critical thinking skills concepts under scrutiny (Supplementary Table 7). These findings indicated that our intervention design based on short educational videos failed to improve participants' understanding and application of the critical thinking concepts being examined in this trial.

Our study was designed to explore the hypothesis that teaching critical thinking skills concepts could enhance fakeness and misinformation recognition skills. However, as discussed above, our intervention did not influence participants' understanding and adoption of critical thinking skills concepts. We sought to test whether our intervention could still improve participants' ability to recognize fakeness and misinformation, regardless of their performance on critical thinking skills concepts. However, the intervention did not impact the participants' ability to identify fakeness (Non-intervention −0.026 versus Intervention −0.025, p = 0.78), misinformation (Non-intervention −0.018 versus Intervention −0.007, p = 0.53), and the combination of fakeness and misinformation (Non-intervention −0.022 versus Intervention −0.016, p = 0.72) (Supplementary Table 8). Therefore, we concluded that our intervention design does not help participants build resilience against fakeness and misinformation.

Given the diverse composition of participants recruited in the study, we tested the hypothesis that independent variables, such as participants' country of origin, age, gender, educational level, survey completion source, and randomization group, might explain the ineffectiveness of the intervention.

However, no single independent variable influenced the effectiveness of our intervention, either in improving the participants' recognition of fakeness and misinformation (Supplementary Table 9) or the participants' understanding and application of critical thinking skills concepts (Supplementary Table 10).

4 Discussion

Results from the pre-intervention phase of our randomized trial, which included a broad range of online participants, suggested that the understanding and correct application of critical thinking skills concepts (i.e., causation and correlation, experimental control, independent data and replicates, statistical significance, reproducibility, and credibility of sources) highly correlate with the proficiency in identifying fakeness and misinformation. Causation and correlation emerged as the most relevant critical thinking skills concepts for unmasking fakeness and misinformation, followed by independent data and replicates, experimental control, reproducibility, statistical significance, and credibility of sources (Figure 2). Such observation may indicate that not all critical thinking skills concepts are equally important for detecting

inaccurate information and may imply a hierarchy of critical thinking skills concepts necessary for effectively distinguishing accurate from inaccurate information. Overall, based on the findings presented in this paper, we argue that the critical thinking skills concepts we operationalized significantly contribute, to varying degrees, in shaping individuals' ability to assess the credibility and accuracy of web-based content. This is particularly relevant in social media and other digital news media environments, where users often need to process an overwhelming volume of readily available and constantly evolving contents. That said, we cannot rule out the possibility that the strong correlation between possessing critical thinking skills and correctly identifying fakeness and misinformation was influenced by the participants' attentiveness – albeit responses from inattentive participants were excluded – and proficiency in answering various survey questions of different complexity and across different topics. If present, this effect could have been amplified by the diverse characteristics of the participants, and in particular the sources through which the participants were recruited (i.e., Facebook, mTurk, and UZH mailing list).
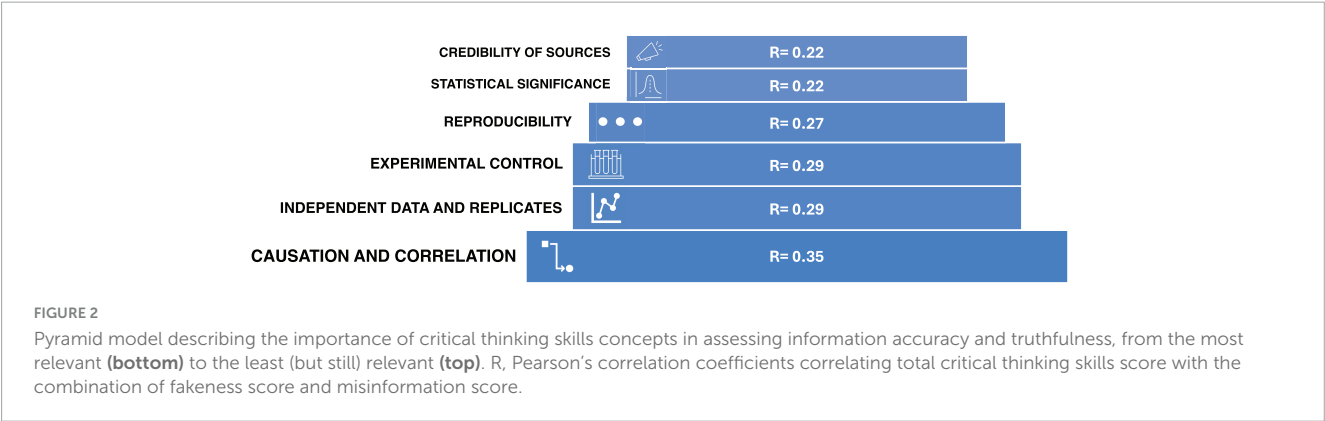
While during the pre-intervention phase of the randomized trial we could show a correlation between possessing critical thinking skills and the ability to recognize fakeness and misinformation, our intervention design failed to effectively train critical thinking skills and/or build resilience against fakeness and misinformation. The various background of participants does not account for this outcome, as independent variables, such as randomization group, country of origin, age, gender, educational level, and survey completion source, did not affect participants' ability to adopt critical thinking skills concepts or detect inaccurate information. One possible explanation for the ineffectiveness of our intervention is that it was not well designed and structured, being inadequate or insufficient for effectively training the critical thinking skills concepts under scrutiny and improving participants' ability to assess the trustworthiness of online information. Alternatively, the intervention may have been cognitively demanding, as it required participants to watch a series of video lectures (see Supplementary Files 1–6), which necessitated sustained mental effort and concentration to be fully absorbed and internalized. Thus, we cannot exclude that the duration of the survey and watching multiple videos consecutively may have impacted the overall effectiveness of the intervention. We should also recognize that educational videos, while proven effective for science education and communication (Brame, 2016; Stockwell et al., 2015), represent only one type of possible intervention. Recent systematic reviews highlight the effectiveness of inquiry-based learning in science education, where students engage in activities such as questioning, experimenting, interpreting data, and constructing evidence-based explanations to deepen their understanding of the natural world (Strat et al., 2024; Urdanivia Alarcon et al., 2023). At the same time, growing attention is directed toward digital educational games and technology-enhanced platforms, such as virtual laboratories – which provide interactive, immersive environments that enhance learning attitudes (Gui et al., 2023; Urdanivia Alarcon et al., 2023).

Our video intervention was not designed to be a customized training that addresses complex and diverse publics, such as those involved in this study; and this limitation is particularly relevant, given that understanding the needs of specific target audiences

TABLE 2 The intervention design is ineffective in training critical thinking skills concepts.

	Group	Mean	Mann-Whitney U	*p*-value
Total critical thinking skills	Non-intervention	0.005	109394.0	ns (0.06)
	Intervention	0.018		

There was no statistically significant difference between the non-intervention and intervention groups in terms of the change in total critical thinking skills scores after the intervention. This was calculated as the difference in means (post-intervention score – pre-intervention score), expressed as normalized delta values: (mean total critical thinking skills score post-intervention – mean total critical thinking skills score pre-intervention)/maximum score. Mann-Whitney U rank test, n(non-intervention) = 334, n(intervention) = 611, *p*-value > 0.05, not significant; post, post intervention; pre, pre intervention.

FIGURE 2
Pyramid model describing the importance of critical thinking skills concepts in assessing information accuracy and truthfulness, from the most relevant **(bottom)** to the least (but still) relevant **(top)**. R, Pearson's correlation coefficients correlating total critical thinking skills score with the combination of fakeness score and misinformation score.

and delivering personalized information is a priority in public health management and in the efforts to address misinformation (Spitale et al., 2024). In addition, our non-intervention group did not receive any form of "placebo" intervention, meaning these participants were not exposed to any videos for the same duration and complexity as the intervention group. This may have created an imbalance between the non-intervention arm and the intervention arm, with non-intervention participants going through the post-phase of the intervention with less cognitive fatigue.

We know that logical and critical reasoning are powerful tools for detecting fake news (Musi and Reed, 2022), countering rhetorical techniques in science denialism (Schmid and Betsch, 2019), and fostering climate literacy (Zanartu et al., 2024). We also know that logical fallacies through inoculation can serve as a valuable tool for recognizing misinformation (Biddlestone et al., 2023; Cook et al., 2017). We also discussed that the inoculation theory has limitations, such as its effectiveness being restricted to the specific content or deceptive strategies it targets, making it unlikely to work as a permanent solution against misinformation. Moreover, only limited evidence has explored the positive role of critical reasoning in detecting deceptive or misleading content – such as the importance of identifying invalid inferences to debunk misinformation (Irving et al., 2022) or evaluating credibility of sources to build resilience against fake news (Soetekouw and Angelopoulos, 2024). Because of these limitations, here we attempted to conceptualize operationalizable, teachable, and trainable critical thinking skills designed to equip the public to autonomously navigate the evolving landscape of misinformation. Building on the literature integrating critical reasoning into scientific literacy (Gormally et al., 2012; Vieira et al., 2011), we realized that critical thinking skills can be seen as the underlying framework of established research practices within the scientific method. Arguably, if key concepts from these practices can be distilled into actionable skills, it is plausible that the public could develop critical reasoning skills that would allow them

to reason like scientists, without the need to absorb academic level knowledge across any field (Redaelli, 2020). However, it is important to emphasize that the limited availability of extensive literature on this topic challenged our ability to determine which critical thinking skills to operationalize and prevented us from designing the intervention based on robust, validated evidence. Due to the pioneering nature of our study, we had to hypothesize the set of measurable critical thinking skills to operationalize in our intervention and determine the most effective approach to training them for improving the recognition of fakeness and misinformation. In the future, we aim to build on our efforts by incorporating additional relevant or more easily teachable critical thinking skills while refining and improving our training approach to develop a more effective intervention design for the concepts operationalized in this study. For example, a gamified intervention may be a suitable way to teach critical thinking for certain segments of the public (Li et al., 2023). That said, it will be crucial to clearly define the target audience and design an intervention tailored to its specific needs. In contrast, the scope of our video intervention in the randomized study may have been overly broad and ambitious, potentially limiting its effectiveness for any of the participant groups involved. We believe it is important to underscore the significance of both positive and negative findings of our study for future research or interventions aimed at improving critical thinking skills: with the findings of this study in hand, we have a clear target; however, we still lack the knowledge on how to effectively implement successful interventions on a large scale.

Critical thinking skills are not only essential for tackling the issue of misinformation and are regarded as one of the ethical aims of effective infodemic management (Germani et al., 2024; World Health Organization (WHO), 2025), but, more broadly, they may also offer a powerful framework within the field of ethics literacy for challenging normative assumptions,

an approach particularly relevant to normatively laden terms like justice, where normative and descriptive dimensions often intrinsically converge (Flores and Burg, 2021). Overall, these observations will strengthen our capacity for everyday ethical and critical reflection on how to evaluate information and reflect upon its content. In line with the results from our study, we expect that the still underexplored landscape of critical thinking skills applied to misinformation recognition will in the future constitute a fundamental toolset to empower different publics through educational approaches, helping them navigate the complexities of the information landscape. We hope this study lays the foundation for identifying the key concepts and skills that should be targeted for training. By doing so, we aim to guide future experimental designs in developing clear and effective interventions.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: original raw data and software used for this study are provided in the study's OSF repository: https://osf.io/f3467/files/osfstorage.

Ethics statement

Ethical approval was not required for the studies involving humans because this study does not fall under the Swiss Human Research Ordinance (HRO) nor the Human Research Act (HRA), and therefore an ethics assessment was not required. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

SR: Conceptualization, Validation, Writing – original draft, Writing – review and editing. NB-A: Methodology, Supervision, Validation, Writing – review and editing. SG: Methodology, Writing – review and editing. JB: Methodology, Writing – review and editing. GS: Investigation, Methodology, Supervision, Validation, Writing – review and editing. FG: Conceptualization, Investigation, Methodology, Supervision, Validation, Visualization, Writing – review and editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that Gen AI was used in the creation of this manuscript. The authors used OpenAI's large language models (LLMs) for editorial assistance.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2025.1577692/full#supplementary-material

References

Andreucci-Annunziata, P., Riedemann, A., Cortes, S., Mellado, A., and del Rio, M. (2023). Vega-Munoz. Conceptualizations and instructional strategies on critical thinking in higher education: A systematic review of systematic reviews. *Front. Educ.* 8:1141686. doi: 10.3389/feduc.2023.1141686

Bates, J., Cheng, S., Ferris, M., and Wang, X. (2025). (Frank). cultivating critical thinking skills: A pedagogical study in a business statistics course. *J. Stat. Data Sci. Educ.* 33, 166–176. doi: 10.1080/26939169.2024.2394534

Bhuttah, T. M., Xusheng, Q., Abid, M. N., and Sharma, S. (2024). Enhancing student critical thinking and learning outcomes through innovative pedagogical approaches in higher education: The mediating role of inclusive leadership. *Sci. Rep.* 14:24362. doi: 10.1038/s41598-024-75379-0

Biddlestone, M., Roozenbeek, J., and Van Der Linden, S. (2023). Once (but not twice) upon a time: Narrative inoculation against conjunction errors indirectly reduces conspiracy beliefs and improves truth discernment. *Appl. Cogn. Psychol.* 37, 304–318. doi: 10.1002/acp.4025

Borges Do Nascimento, I. J., Pizarro, A. B., Almeida, J. M., Azzopardi-Muscat, N., Gonçalves, M. A., et al. (2022). Infodemics and health misinformation: A systematic review of reviews. *Bull. World Health Organ.* 100, 544–561. doi: 10.2471/BLT.21.287654

Brame, C. (2016). Effective educational videos: Principles and guidelines for maximizing student learning from video content. *CBE Life Sci Educ.* 15:es6. doi: 10.1187/cbe.16-03-0125

Cavagnetto, A. R. (2010). Argument to foster scientific literacy: A review of argument interventions in k–12 science contexts. *Rev. Educ. Res.* 80, 336–371. doi: 10.3102/0034654310376953

Cook, J., Lewandowsky, S., and Ecker, U. K. H. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PLoS One* 12:e0175799. doi: 10.1371/journal.pone.0175799

Fitzgerald, M. (1997). Misinformation on the internet: Applying evaluation skills to online information. *Emergency Librarian* 24, 9–14.

Flores, R., and Burg, R. (2021). A case for conscious normativity: Or how ethics literacy can benefit sociology students and their teachers. *Civic Sociol.* 2:18219. doi: 10.1525/CS.2021.18219

Germani, F., Spitale, G., Machiri, S., Ho, C., Ballalai, I., Biller-Andorno, N., et al. (2024). Ethical considerations in infodemic management: Systematic scoping review. *JMIR Infodemiol.* 4:e56307. doi: 10.2196/56307

Gormally, C., Brickman, P., and Lutz, M. (2012). Developing a test of scientific literacy skills (TOSLS): Measuring undergraduates' evaluation of scientific information and arguments. *LSE* 11, 364–377. doi: 10.1187/cbe.12-03-0026

Gui, Y., Cai, Z., Yang, Y., Kong, L., Fan, X., Tai, H., et al. (2023). Effectiveness of digital educational game and game design in STEM learning: A meta-analytic review. *IJ STEM Ed* 10:36. doi: 10.1186/s40594-023-00424-9

Howell, E. L., and Brossard, D. (2021). (Mis)informed about what? What it means to be a science-literate citizen in a digital world. *Proc. Natl. Acad. Sci. U.S.A.* 118, e1912436117. doi: 10.1073/pnas.1912436117

Irving, D., Clark, R. W. A., Lewandowsky, S., and Allen, P. J. (2022). Correcting statistical misinformation about scientific findings in the media: Causation versus correlation. *J. Exp. Psychol. Appl.* 28, 1–9. doi: 10.1037/xap0000408

Lazer, D., Baum, M., Benkler, Y., Berinsky, A., Greenhill, K., Menczer, F., et al. (2018). The science of fake news. *Science* 359, 1094–1096. doi: 10.1126/science.aao2998

Li, M., Ma, S., and Shi, Y. (2023). Examining the effectiveness of gamification as a tool promoting teaching and learning in educational settings: A meta-analysis. *Front. Psychol.* 14:1253549. doi: 10.3389/fpsyg.2023.1253549

Li, Y., Fan, Z., Yuan, X., and Zhang, X. (2022). Recognizing fake information through a developed feature scheme: A user study of health misinformation on social media in China. *Information Process. Manag.* 59:102769. doi: 10.1016/j.ipm.2021.102769

Musi, E., and Reed, C. (2022). From fallacies to semi-fake news: Improving the identification of misinformation triggers across digital media. *Discourse Soc.* 33, 349–370. doi: 10.1177/09579265221076609

National Academies of Sciences, Engineering, and Medicine, Division of Behavioral and Social Sciences and Education, Board on Science Education and Committee on Science Literacy and Public Perception of Science. (2016). *Science Literacy: Concepts, Contexts, and Consequences.* 23595. Washington, DC: National Academies Press, doi: 10.17226/23595

Norris, S. P., and Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Sci. Educ.* 87, 224–240. doi: 10.1002/sce.10066

OECD (2023). *Pisa 2025 Science Framework (draft).* Paris: OECD.

Organización de Naciones Unidas. (2018). *Objetivos de Desarrollo Sostenible.* New York, NY: Organización de Naciones Unidas.

Prokop-Dorner, A., Piłat-Kobla, A., Ślusarczyk, M., Świątkiewicz-Mośny, M., Ożegalska-Łukasik, N., Potysz-Rzyman, A., et al. (2024). Teaching methods for critical thinking in health education of children up to high school: A scoping review. *PLoS One.* 19:e0307094. doi: 10.1371/journal.pone.0307094

Redaelli, S. (2020). *A Guide to Using the Scientific Method in Everyday Life. PLoS SciComm.* Available online at: https://scicomm.plos.org/2020/08/04/a-guide-to-using-the-scientific-method-in-everyday-life/ (accessed April 25, 2025).

Roozenbeek, J., Culloty, E., and Suiter, J. (2023). Countering misinformation: Evidence, knowledge gaps, and implications of current interventions. *Eur. Psychol.* 28, 189–205. doi: 10.1027/1016-9040/a000492

Rubba, P. A., and Andersen, H. O. (1978). Development of an instrument to assess secondary school students understanding of the nature of scientific knowledge. *Sci. Educ.* 62, 449–458. doi: 10.1002/sce.3730620404

Sabzalieva, E., Chacon, E., Bosen, L. L., Morales, D., Mutize, T., Nguyen, H., et al. (2021). *Thinking Higher and Beyond Perspectives on the Futures of Higher Education to 2050.* Paris: UNESCO.

Schmid, P., and Betsch, C. (2019). Effective strategies for rebutting science denialism in public discussions. *Nat. Hum. Behav.* 3, 931–939. doi: 10.1038/s41562-019-0632-4

Shaffer, J. F., Ferguson, J., and Denaro, K. (2019). Use of the test of scientific literacy skills reveals that fundamental literacy is an important contributor to scientific literacy. *LSE* 18:ar31. doi: 10.1187/cbe.18-12-0238

Soetekouw, L., and Angelopoulos, S. (2024). Digital resilience through training protocols: Learning to identify fake news on social media. *Inf. Syst. Front.* 26, 459–475. doi: 10.1007/s10796-021-10240-7

Spitale, G., Germani, F., Biller-Andorno, N., Redaelli, S., Gloeckler, S., and Brown, J. (2022). *osf.io. 2022 can critical thinking skills help us recognize misinformation better? - file storage.* Available online at: https://osf.io/f3467/files/osfstorage (accessed March 17, 2025).

Spitale, G., Germani, F., and Biller-Andorno, N. (2024). The PHERCC matrix. An ethical framework for planning, governing, and evaluating risk and crisis communication in the context of public health emergencies. *Am. J. Bioethics* 24, 67–82. doi: 10.1080/15265161.2023.2201191

Stockwell, B., Stockwell, M., Cennamo, M., and Jiang, E. (2015). Blended learning improves science education. *Cell* 162, 933–936. doi: 10.1016/j.cell.2015.08.009

Strat, T. T. S., Henriksen, E. K., and Jegstad, K. M. (2024). Inquiry-based science education in science teacher education: A systematic review. *Stud. Sci. Educ.* 60, 191–249. doi: 10.1080/03057267.2023.2207148

Tiruneh, D. T., Verburgh, A., and Elen, J. (2014). Effectiveness of critical thinking instruction in higher education: A systematic review of intervention studies. *HES* 4:1. doi: 10.5539/hes.v4n1p1

Urdanivia Alarcon, D. A., Talavera-Mendoza, F., Rucano Paucar, F. H., Cayani Caceres, K. S., and Machaca Viza, R. (2023). Science and inquiry-based teaching and learning: A systematic review. *Front. Educ.* 8:1170487. doi: 10.3389/feduc.2023.1170487

Valladares, L. (2021). Scientific literacy and social transformation: Critical perspectives about science participation and emancipation. *Sci. Educ.* 30, 557–587. doi: 10.1007/s11191-021-00205-2

Vieira, R. M., Tenreiro-Vieira, C., and Martins, I. P. (2011). Critical thinking: Conceptual clarification and its importance in science education. *Critical thinking* 22, 43–54.

Vosoughi, S., Roy, D., and Aral, S. (2018). The spread of true and false news online. *Science* 359, 1146–1151. doi: 10.1126/science.aap9559

Willingham, D. T. (2020). Ask the cognitive scientist: How can educators teach critical thinking? *Am. Educ.* 31, 8–19.

World Health Organization (WHO). (2025). *Social Listening in Infodemic Management for Public Health Emergencies: Guidance on Ethical Considerations.* Geneva: WHO.

World Health Organization (2020). *Immunizing the Public Against Misinformation.* Geneva: WHO.

Yu, L., and Zin, Z. M. (2023). The critical thinking-oriented adaptations of problem-based learning models: A systematic review. *Front. Educ.* 8:1139987. doi: 10.3389/feduc.2023.1139987

Zanartu, F., Cook, J., Wagner, M., and García, J. (2024). A technocognitive approach to detecting fallacies in climate misinformation. *Sci. Rep.* 14:27647. doi: 10.1038/s41598-024-76139-w

Zarocostas, J. (2020). How to fight an infodemic. *Lancet* 395:676. doi: 10.1016/S0140-6736(20)30461-X

Zeng, X., and Ravindran, L. (2025). Design, implementation, and evaluation of peer feedback to develop students' critical thinking: A systematic review from 2010 to 2023. *Thinking Skills Creativity* 55:101691. doi: 10.1016/j.tsc.2024.101691

Zimmerman, T., Njeri, M., Khader, M., and Allen, J. (2022). Default to truth in information behavior: A proposed framework for understanding vulnerability to deceptive information. *ILS* 123, 111–126. doi: 10.1108/ILS-08-2021-0067