Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Spontaneous Sourcing Strategies and Heuristics in Adolescents' Evaluation of Online Information

Margherita Ghiara

m.ghiara@studenti.unisr.it

Università Vita-Salute San Raffaele

Mara Floris

Università Vita-Salute San Raffaele

Martina Calderisi

Università degli Studi di Torino

Yugin Cho

Scuola Universitaria Superiore IUSS Pavia

Carlo Martini

Università Vita-Salute San Raffaele

Article

Keywords: digital literacy, science disinformation, information evaluation, adolescents

Posted Date: March 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-7760447/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

In today's information-rich environment, adolescents face significant challenges in evaluating the credibility of scientific content on social media. This study investigates how 130 high school students assessed the truthfulness of scientific and pseudoscientific Instagram-style posts in a classroom-based, ecologically valid setting. Participants provided written justifications for their judgments, which were thematically coded into 17 evaluative strategies grouped into three macro-categories: lateral reading, cues, and plausibility. Lateral reading was the most frequently used and correlated with higher judgment accuracy, though not significantly. Plausibility-based reasoning, by contrast, often led to incorrect classifications. Over 60% of students conducted external searches, but only half were judged effective. Notably, 40% of participants revised at least one judgment during the activity, and most of these revisions led to improved accuracy. The findings underscore the need for digital literacy interventions that promote reflective reasoning, effective search habits, and critical engagement with diverse epistemic resources.

Introduction

In today's digital environment, adolescents are constantly exposed to an abundance of information, much of which circulates in visually compelling but epistemically ambiguous formats such as social media posts and influencer-generated content [1, 2]. Platforms like Instagram and TikTok play a central role in shaping young people's understanding of the world, including science-related topics that carry both personal and collective implications, like alimentary habits or vaccination behavior. Yet, navigating this information-rich ecosystem requires more than just digital fluency: it demands critical evaluation skills, epistemic vigilance [3], and awareness of the social and technological infrastructures that mediate access to information.

While adolescents are often described as "digital natives" [4], empirical research challenges the assumption that growing up with digital technologies equates to a high level of digital literacy or critical thinking [5]: studies show that young people's ability to assess the credibility of online content is frequently limited by superficial evaluation strategies and cognitive constraints [6, 7]. Rather than applying rigorous analytical frameworks, they tend to rely on visual appeal, personal preference, familiarity, or search engine rankings [1]. Even though such heuristics may be efficient, they are often epistemically unreliable [8, 9].

Indeed, adolescents' evaluative behaviors might be shaped by a combination of factors: bounded rationality, information overload, and satisficing strategies [6, 7, 10]. In practical terms, this means settling for the first seemingly plausible result rather than systematically seeking the most authoritative or corroborated source. This behavior often occurs even when students are aware of traditional credibility markers such as author expertise or institutional affiliation [9, 11]. A consistent finding across observational and experimental studies is the discrepancy between what adolescents claim to do to assess credibility and what they actually do [11, 12].

Adding complexity to the picture is the increasing role of social and contextual cues in shaping how young people evaluate online information. Recent ethnographic research has introduced the notion of information sensibility, i.e. a socially situated awareness of the value of information grounded not only in its epistemic features but also in its relevance to social positioning, identity, and emotional resonance [13, 14]. Adolescents, for instance, may judge credibility based on peer reactions (crowdsourcing) or on the presence of personal, authentic experiences; surrogate trust in influencers or popular personalities; or just superficially explore a topic just to be able to win an argument on social media [14]. These practices challenge normative models of information literacy that assume the primacy of objectivity and evidence.

Despite these challenges, it would equally be incorrect to say that young people are completely gullible because they lack formal training in discriminating between correct and incorrect science-related information. Harris shows that humans have a capacity for evaluating what and whom to believe, a skill observable even in early childhood [15]. One of the most fundamental expressions of this ability is the human tendency to ask questions, a trait that, despite various attempts to cultivate it in other species, remains uniquely human [15, 16]. Questioning serves as a core strategy through which children actively seek out reliable knowledge, combining their own direct observations with information provided by others, in a mix of autonomous and deferential judgement.

As Harris puts it, asking questions is what allows us to "*learn from others*" [15]. Yet this raises an important question: how do we move from the act of questioning to the actual discernment between well-founded and unfounded information? Harris addresses this issue by examining the nuanced balance of autonomy and deference observed in children. He identifies three core strategies that foster the development of epistemic vigilance, i.e., the cognitive faculty that underpins our ability to filter information. These strategies, along with many others, recur across historical periods and disciplinary boundaries [17].

Familiarity with the source is one of such strategies. Children, as well as adults, tend to trust more familiar informants rather than strangers [15], displaying a correlational tendency between epistemic trust and social-emotional closeness that might have evolutionary roots [18]. Such a strategy has also been found among adolescents in online environments [1] and represents one of the main criteria for information evaluation.

However, trust is not given indiscriminately, but is rather moderated by the informants' accuracy. As children grow, they shift from relying primarily on relational familiarity to *tracking the epistemic track record* of informants [19]. This second strategy allows children to overcome familiarity, by trusting those informants who have been accurate in the past. Such a strategy allows for a more sophisticated evaluation, where evidence is weighted more than familiarity.

Finally, a third strategy involves *group consensus* and *majority views*. Several findings show that children are receptive to the presence of consensus among individuals and that such consensus is interpreted as

a trustworthy source of information [15]. This strategy aligns with Hassoun's and colleagues' results showing the impact of peer reactions on credibility assessment of adolescents' [14].

In this study, we record and categorize the search-and-check behavior of high schoolers, partly in order to learn whether the strategies that Harris mentions, for rather simplified informational environments, are also those that we observe in the complex informational environment in which our subjects are checking for reliable information. While the three aforementioned strategies set solid foundations for accurate evaluations of information, they may not be sufficient in our digitalised information ecosystem [20]. As the sophistication of dis/misinformation increases, so does the inadequacy of superficial evaluation strategies. Recent research has highlighted the value of *lateral reading*, a method employed by expert fact-checkers that involves briefly scanning a site before leaving it to explore external sources for contextualization and verification [21]. Lateral reading has been shown to significantly improve credibility judgments and is increasingly recommended as a key component of digital literacy education [21].

Further insights into the effectiveness of evaluation strategies come from studies that compare novices to domain experts. Experts tend to prioritize source information, engage in deeper reliability assessments, and make more selective use of online content [22]. Moreover, intervention programs such as SEEK (Source, Evidence, Explanation, Knowledge) have demonstrated that explicit instruction in evaluative criteria can significantly enhance learners' ability to distinguish between reliable and unreliable sources [23].

Despite these advances, research on adolescents' actual online behavior remains limited, especially in real-world contexts. Much of the existing literature relies on interview data or laboratory tasks, which may not capture the full complexity of an ecologically plausible environment.

To address this gap, the present study adopts an ecological, classroom-based design to examine how adolescents aged 16–19 assess the validity of scientific information they find in Instagram-style posts (i.e. engage in sourcing [24]). Participants were asked to evaluate a series of scientifically valid and invalid posts and to provide written justifications for their judgments. Through an independent thematic analysis of the justifications, we explore the evaluative strategies employed by students and identify the patterns of reasoning underlying both accurate and inaccurate assessments. Our study contributes to the literature by (1) investigating adolescents' critical engagement with science disinformation in an ecological context; and (2) highlighting the interplay of epistemic, social, and affective cues in their reasoning processes.

Results

Through an iterative two-phase thematic analysis of students' written justifications, we developed a comprehensive codebook for the annotation of evaluative strategies employed by adolescents in response to scientific and pseudoscientific online content. The development process followed established qualitative research protocols [25, 26] and involved both inductive and deductive coding by

multiple researchers (see *Methods*). The resulting framework captures a broad spectrum of cognitive, epistemic, and metacognitive strategies and is detailed in the *Supplementary Materials*.

The codebook includes 15 distinct categories, organized into three main macro-categories (i.e., *lateral reading*, *cues*, and *plausibility)*, and two additional meta-codes (i.e., *external search* and *search effectiveness*). Among these, *lateral reading* emerges as particularly frequent and conceptually central across participants' responses. This category stands out in two forms: *source-oriented* and *content-oriented*. In *source-oriented lateral reading*, students evaluate the credibility of a post by seeking external information about the source, such as the author, organization, or platform behind the content. This strategy aligns with findings by Wineburg and McGrew [21], who demonstrated its centrality in expert fact-checking behaviors. By contrast, *content-oriented lateral reading* involves verifying the specific factual claims in a post by consulting other sources. This reflects a distinct level of analytical engagement and correlates with cognitive reflectiveness and resistance to dis/misinformation [27].

Students' justifications of their sourcing behavior were written in Italian and automatically anonymized by the Padlet tool [28] with different pseudonyms. All quotations in this study have been translated by the team and are followed by the initials of the relative pseudonym assigned by Padlet. Among the *lateral reading* macro-category, a common sub-strategy is *consensus search*, whereby students assess the convergence of information across multiple sources (e.g., "*After examining various sources regarding this news, searching the web, and carefully observing various official newspaper websites, we can confirm that this news is, in my opinion, true. In conclusion, the facts reported in this article are true*", A.F.). This triangulation method is supported by work in social epistemology and civic online reasoning [29, 30]. Closely related is the use of *scientific article verification*, which involves citing or referencing peer-reviewed sources or scientific summaries, a behavior previously found to correlate with epistemic trust in science and higher-level digital literacy [31]. Below is an example of the latter:

"*The study conducted by the University of California, San Diego, published in Archives of Internal Medicine (*https://www.ncbi.nlm.nih.gov/pubmed/20421555*), does not confirm the thesis reported in any way. In fact, the research analyzes a possible link between certain factors and mood, but does not provide any scientific evidence to support the correlation stated in the news*", D.M..

Moreover, a smaller but significant number of students referenced fact-checking or science communication websites like *Facta*, *Focus* or *Geopop*, which are relatively well-known in Italy, demonstrating awareness of debunking infrastructures [32]. Others used *ChatGPT* or similar AI tools for verification, sometimes explicitly and sometimes inferred through stylistic clues [33]. Given the growing use of generative AI in youth information behavior, this strategy warrants particular attention:

"*I don't have any specific sources from which I drew my ideas, but following a discussion with ChatGpt and based on my knowledge from the world of fitness, I came to the conclusion that dark chocolate is a source of fat and therefore makes you gain weight*", H.O..

Some participants relied on *Wikipedia*, a tool often treated with skepticism in academic settings but shown to offer relatively reliable information for general topics [34]. Similarly, *YouTube* and other video-based sources were mentioned, indicating a shift toward audiovisual verification strategies, especially common among adolescents [35]. Some students, moreover, combined the search tools used, like B.D., who provided a thorough explanation of why he believed the post to be true, citing both Wikipedia and Youtube as sources.

Cues strategies encompass heuristic-based approaches such as *source reliability judgments* (e.g., "*According to me, Geopop is always right*", E.F.), *language as a cue* (e.g., "*I decided to read the article carefully, and after reading it, I realized that it was written in a manner consistent with the language of a scientist [...]*", C.C.), and *image-based validation* (e.g., "*In my opinion, it's true because there are also photos*", A.B.), all of which echo well-documented cognitive shortcuts in digital contexts [36, 37].

Other strategies, categorized within the *plausibility* macro-category, center on the *content's internal coherence (*e.g., "*Reading the entire article, it is clear that the news is true*", B.C.), or on plausibility assessments based on either *general reasoning* (e.g., "*In my opinion, it is not possible because you cannot revive a dead animal*", M.P.) or *specific prior knowledge* (e.g., "*As a competitive athlete at the national level, I am fairly knowledgeable about the benefits of various foods, and dark chocolate, which has a high cocoa content, also has a high protein content that promotes muscle growth and fat loss*", H.P.). These plausibility-based judgments highlight the tension between intuitive thinking and formal verification [38, 39].

Finally, we included two meta-codes: *external search* (whether a search was performed and declared) and *search effectiveness*, indicating whether the search led to an accurate judgment supported by evidence [9, 40]. Performing external searches, indeed, does not necessarily lead to proper assessments of information pieces. An example of that was the case of C.H.: When reading a news item stating that a Bachelor's student contracted HIV in a laboratory while carrying on an internship, C.H. didn't verify the truthfulness of the assertion, but only referenced the webpage of the Italian Health Ministry with information related to the means of transmission of the virus.

The 174 student responses were coded using the finalized codebook by two independent annotators, with high inter-rater agreements (Cohen's $\kappa = 0.84$ for lateral reading, $\kappa = 0.78$ for cues, $\kappa = 0.64$ for plausibility, $\kappa = 0.82$ for *external search* and $\kappa = 0.74$ for *search effectiveness*), demonstrating reliable identification of these evaluative strategies [41].

Across all responses, approximately two out of three evaluations (i.e., 64.25%) are correct, confirming a moderate level of discriminatory ability. Incidentally, this rate of accuracy is consistent with previous work by Martini et al. on high-school student ability to discriminate between science-based information and science disinformation [42]. Specifically, 69.4% of students who used a computer and 61.8% of those using a smartphone made correct judgments, though this difference was not statistically significant ($\chi^2 = 0.78$, df = 1, $p = 0.378$).

A frequency analysis revealed notable differences in the adoption of evaluative strategies among students. Lateral reading strategies (including sub-strategies: consensus search, fact-checking websites, generative AI tools, Wikipedia, YouTube, and scientific article verification) are the most prevalent, representing approximately 61% of annotations for Rater 1 and 57% for Rater 2. Among the lateral reading strategies, however, some are under-represented (Fig. 1). In particular, very few students applied lateral reading strategies to compare the content with available scientific literature (2.68% for Rater 1 and 6.25% for Rater 2) or to gain more information about the source of the news item (5.36% for Rater 1 and 7.81% for Rater 2).

Plausibility assessments (*general plausibility*, *specific plausibility*, *internal coherence*, and other intuitive evaluations) account for around 29% of annotations for Rater 1 and 33% for Rater 2. Finally, *cues* strategies (*misleading titles*, *source reliability judgments*, *visual cues*, and *language cues*) are employed less frequently, accounting for approximately 10% of annotations for Rater 1 and 9% for Rater 2. These frequencies indicate clear patterns in adolescents' verification behaviors and preferences, as well as an overall level of critical engagement with the information they were given.

More specifically, Table 1 summarizes the mean frequency of each strategy between the two raters, disaggregated by the six posts presented in the activity. The posts included three scientifically valid and three pseudoscientific items, each designed to mimic the stylistic features of real social media content (See *Methods*).

Table 1
The table displays the relative frequency (%) of each evaluative strategy (i.e., cues, lateral reading, and plausibility) disaggregated by news item. Lateral reading was the most common strategy across all posts, followed by plausibility-reasoning and cue-based strategies.

News item	Cues (%)	Lateral reading (%)	Plausibility (%)
Flamingos (Valid)	10.2	57.8	32.0
HIV (Valid)	5.4	57.7	36.9
Spiders (Valid)	16.7	55.9	27.5
Chocolate (Invalid)	10.9	47.5	41.6
Pyramid (Invalid)	7.8	70.7	21.6
Spain flood (Invalid)	6.1	76.8	17.1

Regarding frequencies of *external search* behavior, 33.7% of times no external search was performed when evaluating the credibility of the posts, whereas in 34.0% and 25.6% of cases multiple sources or a single source were respectively consulted. Finally, in 6.7% cases, students declared to have conducted an external search without specifying the sources consulted. Importantly, in 52.1% cases, the external search was rated as effective (i.e., a search that contributes to an accurate judgment based on meaningful evidence), a result showing that almost half of the external searches were not effective.

Accuracy rates vary substantially depending on the strategy employed. *Lateral reading* is associated with higher accuracy: cases where students used it led to correct evaluations in 71–72% of cases, compared to 58% of accurate evaluations when the strategy was not used. However, such a difference did not result to be significant (p_{rater1} = 0.057, p_{rater2} = 0.086). *Cues*-based reasoning shows accuracy levels between 64–76%, without significant differences between usage or non-usage (p_{rater1} = 0.332, p_{rater2} = 0.155), while *plausibility*-based strategies led to decreased accuracy, especially among Rater 2, where it significantly dropped to 55% when used, compared to 73% when not used (p_{rater1} = 0.494, p_{rater2} = 0.0198).

Finally, analysis of post-activity responses revealed meaningful revisions in students' thinking. In cases when students initially judged a post as true, 18.4% later reported having changed their mind to 'false' after searching, and 70.3% of these revised answers were correct. Conversely, among the cases where students initially judged a post as false, 16.5% changed their mind to 'true', with a correctness rate of 52.4%. Those who maintained their initial judgment had slightly lower but still meaningful accuracy levels (63.4% and 81.1%, respectively, for true and false initial responses).

Discussion

This study investigated how adolescents evaluate the credibility of scientific and pseudoscientific content encountered in Instagram-style posts in a real-world setting. By combining a digital simulation of social media environments with a qualitative and quantitative analysis of students' justifications for their sourcing strategies, we identified a wide range of epistemic strategies employed by young people to assess the validity of information online.

Our findings highlight several notable trends. Firstly, a significant proportion of participants effectively employed advanced evaluative strategies, such as *lateral reading* (primarily *consensus search* and *debunking search)* and specific cues, such as *trustworthy external sources*. Indeed, *lateral reading* emerged as the most frequently employed strategy, positively correlating with higher accuracy in judgements, although this correlation was not significant. This is a promising result, as lateral reading strategies have been proven to be an effective evaluation approach [30]. However, it is important to note that students primarily verified information related to the content. Only a small number of cases involved the application of experts' evaluation strategies, such as *source-oriented lateral reading* or *scientific article verification* [22].

Importantly, the widespread reliance on *content-oriented lateral reading* that was observed in this study may (at least partly) be due to the instructions provided on Padlet boards by the experimenters. Such instructions (which included the explicit invitation to navigate the web, collect some pieces of information and upload them on Padlet) could have affected students' evaluative strategy selection. Hence, we submit, this potential priming effect on adolescents' choice of verification methods should be carefully addressed and measured in future studies.

At the same time, a substantial subset of students continued to rely on *plausibility*-based heuristics, intuition, and *cues* derived from familiarity, language, or visual presentation. These results align with previous literature highlighting adolescents' tendency toward intuitive reasoning and heuristic shortcuts in online contexts [6, 9, 39], as well as with findings from the field of developmental psychology [15], supporting the idea that some of these strategies are embedded in human cognition as some of the work in Harris shows [15].

The study identified hybrid evaluative behaviors as well, reflecting students' pragmatic engagement with available digital tools and resources. A notable portion of students integrated contemporary digital tools, such as generative AI (i.e., ChatGPT), video platforms (i.e., YouTube), and online encyclopedia (i.e., Wikipedia), into their verification processes.

The increasing integration of generative AI and polished digital content underscores the risk of reinforcing the perceived validity of pseudoscientific claims, an issue highlighted in recent discussions about algorithmic trust and cognitive biases [33, 37]. Given the widespread usage of generative AI as a tool to evaluate information, it is pivotal to understand adolescents' awareness of the reliability of information provided by such tools and how they engage with them in order to develop tailored digital interventions.

Moreover, in only half of the cases where an external search was performed, such a search led to an accurate judgment based on meaningful evidence. Taken together, these results highlight the need for digital literacy interventions that go beyond an uncritical application of external search, but rather aim at providing more nuanced tools for distinguishing between a case of information based on science from a case of information based on the mere appearance of science (i.e. pseudoscience).

In real-world contexts, valid information and invalid information traits can be found simultaneously, making it difficult to evaluate content reliability and accuracy. This was the case for some of the information pieces used in the present study, where, for example, a misleading title stating that dead spiders have been revived was attached to a scientifically valid news or where fragments of accurate information were merged to argue in favour of a pseudoscientific claim according to which dark chocolate has been shown to cause weight loss. Educational interventions that integrate multiple strategies together should thus be developed. Such interventions would enable not only to overcome this issue, but also to avoid an uncritical application of effective techniques, such as the application of content-oriented lateral reading to exclusively confirm pre-existing beliefs.

Furthermore, the ecological and classroom-based approach of our study provided insights into students' evaluation strategies within authentic, socially embedded contexts. The use of collaborative yet individually moderated platforms such as Notion and Padlet mirrored realistic digital interactions. This design allowed us to observe the complex interplay between individual critical thinking, social influences, and environmental factors influencing students' evaluative behaviors, providing valuable insights that go beyond those typically obtained in controlled laboratory settings.

However, it should be pointed out that the classroom environment in which our study was carried out differed from real-world scenarios in some important respects. Indeed, students' level of engagement and attention might be expected to be higher at school when they are involved in educational initiatives under the supervision of their teachers than while doing everyday activities out of school. Moreover, students were given a significant amount of time (up to one hour) to complete the requested task, namely assessing the accuracy of at least one piece of news. In real life, adolescents rarely devote a comparable amount of time to the verification of online contents, with the plausible exception of the cases in which these are related to issues they particularly care about.

The effect of the amount of time spent reading and evaluating online news on the accuracy of information assessment should be further investigated in future research, being properly disentangled from the evaluation strategy used. This could be useful to identify the strategies that ensure an acceptable degree of accuracy while also being compatible with time constraints typical of real-life contexts.

The effect of adolescents' motivation on the accuracy of their credibility judgments should be examined in future work, too. By "motivation", we mean one's interest in the topic covered by a certain piece of information (to be controlled for by directly asking participants to rate it before assessing the veracity of some news). We think that motivation (in the sense of the word just specified, not to be confused with motivated reasoning which applies to ethically non-neutral information) is positively correlated with evaluation accuracy. The existence of this correlation should be tested in further studies.

The results of our study suggests that the kind of device (computer vs. smartphone) used to search for relevant information on the internet could have some impact on the accuracy of assessments made by young people. In particular, 69.4% of students who used a computer provided accurate judgments, compared to 61.8% of those using a smartphone, though this gap in accuracy turned out to be not statistically significant. The effect of the specific tool through which information evaluation is carried out should be more systematically addressed in the future, given the increasing centrality of smartphones in teenagers' (as well as adults') lives.

From an educational perspective, our findings emphasize the need for comprehensive digital literacy interventions that go further than checklist-based approaches. Effective educational curriculum should prioritize teaching lateral reading, nuanced source analysis, and epistemic humility; skills that significantly enhance evaluative accuracy yet are rarely intuitive without explicit training [21, 30]. Moreover, given the evolving digital landscape, characterized by generative AI, influencer culture, and interactive media, digital literacy curricula should incorporate critical reflection on the broader infrastructure underpinning knowledge creation and dissemination, as well as providing useful guidelines for AI usage when applying fact-checking strategies.

Lastly, the annotated dataset and comprehensive codebook generated from this study constitute valuable resources for future educational research. We believe that they offer empirical foundations for developing pedagogical tools specifically tailored to adolescent learners. Yet, the lower inter-rater

agreement result for the category *plausibility* with respect to the other categories presented and discussed here represents an additional limit of the present study, that should be overcome by future research through a clearer category definition and/or through appropriate rater training.

Methods

Participants

Data collection was conducted between May and June 2025 in two upper secondary schools located in the metropolitan area of a large city in northern Italy. The schools were recruited through a public engagement initiative led by a local university: invitations were sent to publicly available institutional email addresses starting on 01/09/2024, and the first two schools to respond were selected, with participant recruitment ending on 31/10/2024. A total of 130 students from multiple fourth-grade classes (aged 16–19) participated in the study. All students in the selected classes were invited to participate, with nobody choosing to opt out. No exclusion criteria were applied aside from age and school year, and no demographic information was collected.

The activity took place during regular school hours and was supervised by members of the research team in collaboration with classroom teachers. Participation was entirely voluntary, and students could withdraw at any time without penalty. Written informed consent was obtained from all participants prior to data collection through paper-based consent forms signed by each student. In compliance with the General Data Protection Regulation (Regulation EU 2016/679), no personally identifiable information was collected. Ethical approval for the study was provided by the Lombardy Regional Ethics Committee 1 (approval number CET 175–2025). Additionally, all research was performed in accordance with relevant guidelines and regulations, as well as in accordance with the Declaration of Helsinki. The Padlet tool used in the study does not store IP addresses or user metadata, and all student responses were anonymized at the point of data collection.

Study Procedure

Participants took part in a classroom-based session of approximately two hours, facilitated by a member of the research team. The session began with a 30-minute presentation introducing the research team, the overall study, and key concepts such as misinformation, disinformation, and their implications in scientific contexts.

Following the introduction, students were directed to a custom digital environment built on the Notion platform [43], which featured six Instagram-styled posts (see *Supplementary material*s). All posts were based on news pieces the experimenters found on the web, and were selected and independently evaluated by pairs of researchers from the research team. Half of the posts were based on pseudoscientific claims and contained information that was shown to be unequivocally false or misleading by professional fact-checking organisations. The remaining half of the posts contained scientifically accurate claims supported by published scientific evidence. All the themes were selected

based on scientific topics with neutral political resonance in order to avoid interference in evaluations with political motivated reasoning.

Scientifically valid posts included statements about (i) the reasons behind the colour pink of flamingos, (ii) an HIV infection happened in a laboratory, and (iii) the results of a research where dead spiders were turned into ready-to-use actuators. Instead, invalid and pseudoscientific posts included statements about (i) weight-loss properties of dark chocolate, (ii) new archaeological results discussing the actual oldest pyramid known until now, and (iii) a cause-effect relationship between geoengineering and recent floods in Spain. Each post was linked to its original online source and accompanied by two emoji-buttons (i.e., □ and □) that redirected students to specific Padlet boards.

Given the high interactivity design of Padlet, this free online tool has been widely used for educational purposes. Research has previously highlighted how the tool motivates students to participate, exchange information and express opinions comfortably, while fostering effective dialogues by allowing students to learn from peers' responses [44, 45]. In particular, the Padlet tool allows the creation of digital boards, where students are able to anonymously share various media formats, including text, image, audio, video and external URLs. The creators of Padlet boards can control the content, design, privacy settings and access to the board.

Two Padlet boards were therefore created for each news piece, one for the students who evaluated the post as true and one for those who evaluated it to be false (Fig. 3). This design allowed us to manage in the best possible way the debriefing phase, as facilitators could easily pick a few news items to focus on, usually the ones mostly analyzed by students in the classroom, and walk through the different boards to discuss the rationales underlying the responses given, on the one side, by participants who believed the post to be true and, on the other side, by those who believed the post to be false.

Once landed to the relative Padlet digital board, students were asked to:

1. Judge whether the content of each post was true or false.
2. Provide a brief written justification explaining their reasoning. This could include written statements, comments, links, and audio-visual material.
3. Answer an anonymous poll asking whether they still believed the post to be true or false, or whether the activity led them to revise their evaluations.

Participants were encouraged to evaluate at least one of the six posts, although some students chose to assess multiple items. As a result, the dataset comprises 174 individual responses. Each student was allowed to use either their personal smartphone or a school-provided device (i.e., laptop or desktop) to complete the activity. They had between 45 and 60 minutes to formulate their responses, during which they were permitted to consult the internet.

To minimize peer influence, students' sourcing strategies and justifications were not made immediately visible: all content had to be approved by the facilitator before appearing on the shared Padlet board.

During a short 15-minute break, the facilitator reviewed and approved submissions. The session concluded with a debriefing phase, during which the nature of the posts and students' rationales were discussed collectively by walking through students' comments while fostering dialogue: facilitators provided key take-aways on digital critical thinking and effective ways to evaluate news pieces, and students were free to discuss with the facilitators and the rest of the class. The Padlet tool enabled the collection of both categorical (true/false) and open-ended responses in a fully anonymous format.

Thematic Analysis and Coding Scheme

The qualitative analysis of student justifications followed a two-phase thematic approach, combining both inductive and deductive elements [25, 26]. The aim was to identify and classify the evaluative strategies adolescents used to assess the validity of scientific and pseudoscientific Instagram-style posts.

In the first phase, three members of the research team (C.M., M.G., and M.F.) independently reviewed the full set of 174 open-ended responses. Each researcher generated initial codes based on recurring patterns in the data, focusing on the epistemic, cognitive, and social cues employed by students. This initial round of open coding produced a diverse set of labels, which were then compared and refined through an iterative process of discussion and negotiated consensus. Through successive rounds of synthesis, overlapping codes were merged, hierarchical relationships were established (e.g., main categories and subcategories), and ambiguous cases were clarified. The result was the development of a structured codebook, comprising 17 coding labels that represent distinct credibility assessment strategies. The full codebook, including definitions, examples, and theoretical references, is available in the *Supplementary Information*.

In the second phase, the finalized codebook was applied to the entire dataset by two independent annotators (M.C. and Y.C.), who had not participated in the initial code development. Prior to annotation, both coders received a brief training session on the use of the codebook. Each student response could receive multiple labels if more than one strategy was present. To assess the reliability of the coding scheme, inter-rater agreement was calculated using Cohen's Kappa, a widely adopted measure of categorical consistency. The analysis yielded moderate to high levels of agreement ($\kappa = 0.84$ for *lateral reading*, $\kappa = 0.78$ for *cues*, $\kappa = 0.64$ for *plausibility*, $\kappa = 0.82$ for *external search* and $\kappa = 0.74$ for *search effectiveness*), indicating that the codebook was both conceptually clear and reproducible across raters.

The annotated dataset allowed for a quantitative analysis of the frequency and distribution of credibility strategies across responses and post types. In addition to supporting the present study, this structured dataset can serve as a resource for future research on adolescent reasoning and for the development of educational and computational tools.

Data Analysis

Data were analyzed through a mixed-method approach [46] in three complementary stages: frequency-based exploration of evaluative strategies, assessment of judgment accuracy, and examination of search behavior and revision patterns. All analyses were performed using R Statistical Software [47].

In the first stage, we computed the frequency of each strategy identified in the thematic coding, disaggregated by rater and post type. Each of the 174 student responses could include multiple strategies, and all coded strategies were included in the analysis. For comparative purposes, strategies were grouped into three macro-categories: *lateral reading, cues*, and *plausibility*. Sub-strategies (e.g., *use of generative AI tool, Wikipedia, or scientific article validation*) were nested within the relevant macro-category.

In the second stage, we assessed the accuracy of students' true/false judgments, defined as alignment between the student's judgment (true/false) and the actual validity of the information. Accuracy was calculated both globally and in relation to strategy use. For each strategy, we compared the accuracy of responses where the strategy was used to those where it was not. This enabled us to identify associations between evaluative behavior and outcome quality and their significance through t-test statistics. Lateral reading, in particular, was associated with higher accuracy levels, although the difference was not significant ($t_{rater1} = 1.92$, $df_{rater1} = 133.64$; $p_{rater1} = 0.057$; $t_{rater2} = 1.73$, $df_{rater2} = 114.8$, $p_{rater2} = 0.086$). Cues-based reasoning resulted in accuracy levels between 64–76%, without significant differences between usage or non-usage ($t_{rater1} = 0.98$, $df_{rater1} = 52.04$; $p_{rater1} = 0.332$; $t_{rater2} = 1.44$, $df_{rater2} = 53.89$, $p_{rater2} = 0.155$), whereas reliance on plausibility reasoning often coincided with lower performance ($t_{rater1} = -0.69$, $df_{rater1} = 88.29$; $p_{rater1} = 0.494$; $t_{rater2} = -2.36$, $df_{rater2} = 120.53$, $p_{rater2} = 0.02$).

The third stage focused on students' search behavior and its effectiveness. Each justification was coded for whether a search was performed and, if so, whether it contributed meaningfully to the accuracy of the final judgment. Frequencies of search use were categorized as no search, single-source search, multi-source search, or unspecified search. Search effectiveness was rated qualitatively based on whether the participant's search led to an accurate conclusion inferred from relevant pieces of information. Inter-rater agreement was high for both dimensions ($\kappa = 0.82$ for external search; $\kappa = 0.74$ for search effectiveness). Among those cases when searches were conducted, just over half (i.e., 52.1%) were judged to be effective.

Finally, we analyzed revisions in students' judgments across the activity. Some students initially evaluated a post, then modified their judgment after engaging with further information. By comparing initial and final responses (where available), we documented shifts in opinion and calculated the proportion of judgment changes that resulted in a correct response. Approximately 40% of students changed at least one answer after reflection or search, and these changes led to improved accuracy in the majority of cases. This finding highlights the potential of encouraging reflective digital practices, particularly when combined with explicit training in evaluative strategies.

Declarations

Data availability

The raw dataset, labelled datasets, and materials related to this work have been deposited in the Open Science Framework (OSF) repositories for this project and are available at the following link https://osf.io/gn3br/?view_only=888f8aa5f9734231ba24c478279346f6.

Additional Information

The authors declare no competing interests.

Funding Declaration

The study was supported by the European Union – Next Generation EU, Mission 4, Component 1 (CUP D46F23000120004).

References

1. Freeman, J. L., Caldwell, P. H. Y. & Scott, K. M. How Adolescents Trust Health Information on Social Media: A Systematic Review. *Acad. Pediatr.* **23,** 703–719 (2023).
2. Boczkowski, P. J. *Abundance: On the Experience of Living in a World of Information Plenty*. (Oxford University Press, 2021).
3. Brown, P. & Gummerum, M. Trust issues: Adolescents' epistemic vigilance towards online sources. *Br. J. Dev. Psychol.* (2025). doi:10.1111/bjdp.12559
4. Prensky, M. Digital Natives, Digital Immigrants Part 1. *Horiz.* **9,** 1–6 (2001).
5. Helsper, E. J. & Eynon, R. Digital natives: Where is the evidence? *Br. Educ. Res. J.* **36,** 503–520 (2010).
6. Agosto, D. E. A model of young people's decision-making in using the Web. *Libr. Inf. Sci. Res.* **24,** 311–341 (2002).
7. Gasser, U., Cortesi, S., Malik, M. M. & Lee, A. Youth and Digital Media: From Credibility to Information Quality. SSRN Scholarly Paper at https://doi.org/10.2139/ssrn.2005272 (2012)
8. Pan, B. *et al.* In Google We Trust: Users' Decisions on Rank, Position, and Relevance. *J. Comput.-Mediat. Commun.* **12,** 801–823 (2007).
9. Hargittai, E., Fullerton, L., Menchen-Trevino, E. & Thomas, K. Y. Trust Online: Young Adults' Evaluation of Web Content. *Int. J. Commun.* **4,** 27 (2010).
10. McGrew, S. & Kohnen, A. M. Tackling misinformation through online information literacy: Structural and contextual considerations. *J. Res. Technol. Educ.* **56,** 1–6 (2024).
11. Iding, M. K., Crosby, M. E., Auernheimer, B. & Klemm, E. B. Web site credibility: Why do people believe what they believe? *Instr. Sci.* **37,** 43–63 (2009).
12. Flanagin, A. J. & Metzger, M. J. The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media Soc.* **9,** 319–342

(2007).

13. Hassoun, A. *et al.* Practicing Information Sensibility: How Gen Z Engages with Online Information. in *Proc. 2023 CHI Conf. Hum. Factors Comput. Syst.* 1–17 (ACM, 2023). doi:10.1145/3544548.3581328
14. Hassoun, A. *et al.* Beyond Digital Literacy: Building Youth Digital Resilience Through Existing "Information Sensibility" Practices. *Soc. Sci.* **14,** 230 (2025).
15. Harris, P. L. *Trusting what you're told: How children learn from others*. 253 (The Belknap Press of Harvard University Press, 2012).
16. Greenfield, P. M. & Savage-Rumbaugh, E. S. Comparing communicative competence in child and chimp: the pragmatics of repetition. *J. Child Lang.* **20,** 1–26 (1993).
17. Martini, C. in *Routledge Handb. Soc. Epistemol.* (Routledge, 2019).
18. Mercier, H. *Not Born Yesterday: The Science of Who We Trust and What We Believe*. (Princeton University Press, 2020). doi:10.1515/9780691198842
19. Corriveau, K. & Harris, P. L. Choosing your informant: weighing familiarity and recent accuracy. *Dev. Sci.* **12,** 426–437 (2009).
20. Wineburg, S. & McGrew, S. Why Students Can't Google Their Way to the Truth. *Educ. Week* (2016). at <https://www.edweek.org/teaching-learning/opinion-why-students-cant-google-their-way-to-the-truth/2016/11>
21. Wineburg, S. & McGrew, S. Lateral Reading: Reading Less and Learning More When Evaluating Digital Information. *SSRN Electron. J.* (2017). doi:10.2139/ssrn.3048994
22. Brand-Gruwel, S., Kammerer, Y., Van Meeuwen, L. & Van Gog, T. Source evaluation of domain experts and novices during Web search. *J. Comput. Assist. Learn.* **33,** 234–251 (2017).
23. Wiley, J. *et al.* Source Evaluation, Comprehension, and Learning in Internet Science Inquiry Tasks. *Am. Educ. Res. J.* **46,** 1060–1106 (2009).
24. Brante, E. W. & Strømsø, H. I. Sourcing in text comprehension: A review of interventions targeting sourcing skills. *Educ. Psychol. Rev.* **30,** 773–799 (2018).
25. Braun, V. & Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **3,** 77–101 (2006).
26. Braun, V. & Clarke, V. Reflecting on reflexive thematic analysis. *Qual. Res. Sport Exerc. Health* **11,** 589–597 (2019).
27. Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G. & Cannon, T. D. Belief in Fake News is Associated with Delusionality, Dogmatism, Religious Fundamentalism, and Reduced Analytic Thinking. *J. Appl. Res. Mem. Cogn.* **8,** 108–117 (2019).
28. Padlet - Visual Collaboration for Creative Work and Education. *Padlet* at <https://padlet.com/>
29. Goldman, A. I. Experts: Which Ones Should You Trust? *Philos. Phenomenol. Res.* **63,** 85–110 (2001).
30. McGrew, S., Breakstone, J., Ortega, T., Smith, M. & Wineburg, S. Can Students Evaluate Online Sources? Learning From Assessments of Civic Online Reasoning. *Theory Res. Soc. Educ.* **46,** 165–193 (2018).

31. Barzilai, S. & Zohar, A. Epistemic Thinking in Action: Evaluating and Integrating Online Sources. *Cogn. Instr.* **30,** 39–85 (2012).
32. Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G. & Rand, D. G. Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychol. Sci.* **31,** 770–780 (2020).
33. Bender, E. M., Gebru, T., McMillan-Major, A. & Shmitchell, S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? . in *Proc. 2021 ACM Conf. Fairness Account. Transpar.* 610–623 (Association for Computing Machinery, 2021). doi:10.1145/3442188.3445922
34. Jemielniak, D. Wikipedia, a professor's best friend. *Chron. High. Educ.* (2014).
35. Swart, J. & Broersma, M. The Trust Gap: Young People's Tactics for Assessing the Reliability of Political News. *Int. J. Press.* **27,** 396–416 (2022).
36. Lewandowsky, S., Ecker, U. K. H. & Cook, J. Beyond Misinformation: Understanding and Coping with the "Post-Truth" Era. *J. Appl. Res. Mem. Cogn.* **6,** 353–369 (2017).
37. Newman, E. J., Garry, M., Bernstein, D. M., Kantner, J. & Lindsay, D. S. Nonprobative photographs (or words) inflate truthiness. *Psychon. Bull. Rev.* **19,** 969–974 (2012).
38. Nickerson, R. S. Confirmation bias: A ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* **2,** 175–220 (1998).
39. Kahneman, D. *Thinking, fast and slow*. (2011).
40. Pirolli, P. & Card, S. Information foraging. *Psychol. Rev.* **106,** 643–675 (1999).
41. Landis, J. R. & Koch, G. G. The measurement of observer agreement for categorical data. *Biometrics* **33,** 159–174 (1977).
42. Martini, C., Floris, M. & Ghiara, M. *Disinformazione a Scuola: il progetto UniSR per sviluppare le capacità critiche digitali*. (2024). at <https://www.unisr.it/news/2024/10/disinformazione-scuole-ricerca-carlo-martini>
43. Notion Labs Inc. at <https://www.notion.so/>
44. Arouri, Y. M., Hamaidi, D. A., Al-Kaabi, A. F., Al Attiyah, A. A. & ElKhouly, M. M. Undergraduate Students' Perceptions on the Use of Padlet as an Educational Tool for an Academic Engagement: Qualitative Study. | EBSCOhost. **18,** 86 (2023).
45. Deni, A. R. M. & Zainal, Z. I. Padlet as an educational tool: pedagogical considerations and lessons learnt. in *Proc. 10th Int. Conf. Educ. Technol. Comput.* 156–162 (Association for Computing Machinery, 2018). doi:10.1145/3290511.3290512
46. Creswell, J. W. in *Handb. Educ. Policy* (ed. Cizek, G. J.) 455–472 (Academic Press, 1999). doi:10.1016/B978-012174698-8/50045-X
47. R Core Team. R: A Language and Environment for Statistical Computing. (2025). at <https://www.R-project.org/>

Figures

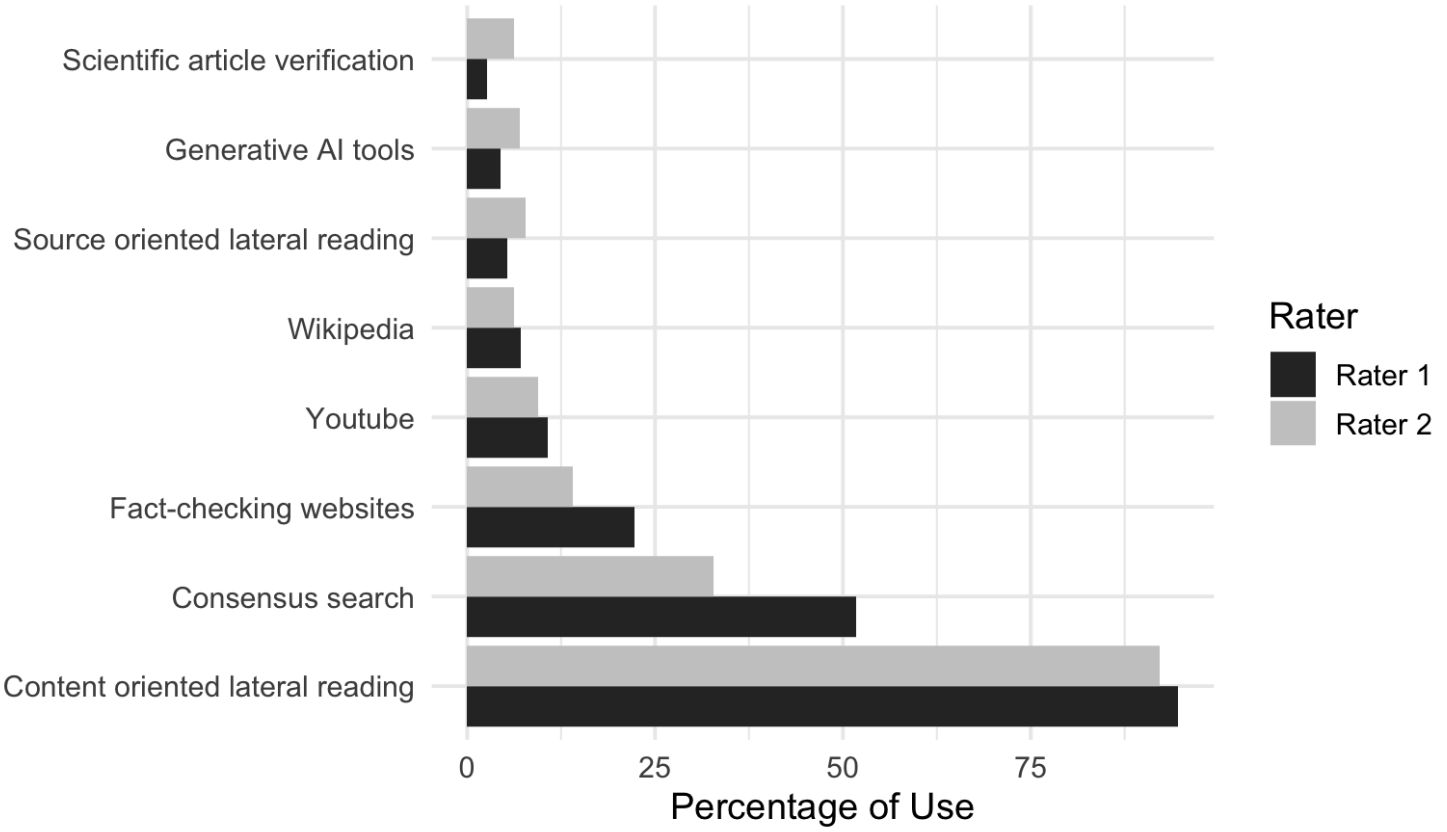


Figure 1

Lateral reading sub-strategies by rater. The plot shows the percentage use of lateral reading sub-strategies coded by two raters. Content-oriented lateral reading was dominant (94.64% for Rater 1 and 92.19% for Rater 2), followed by consensus search (51.79% for Rater 1 and 32.81% for Rater 2) and fact-checking websites (22.32% for Rater 1 and 14.06% for Rater 2). Less frequent was Source-oriented lateral reading (5.36% for Rater 1 and 7.81% for Rater 2), as well as Wikipedia (7.14% for Rater 1 and 6.25% for Rater 2), YouTube (10.71% for Rater 1 and 9.38% for Rater 2), generative AI tools (4.46% for Rater 1 and 7.03% for Rater 2), and scientific article verification (2.68% for Rater 1 and 6.25% for Rater 2). The pattern suggests students favored verifying claims over assessing source credibility.

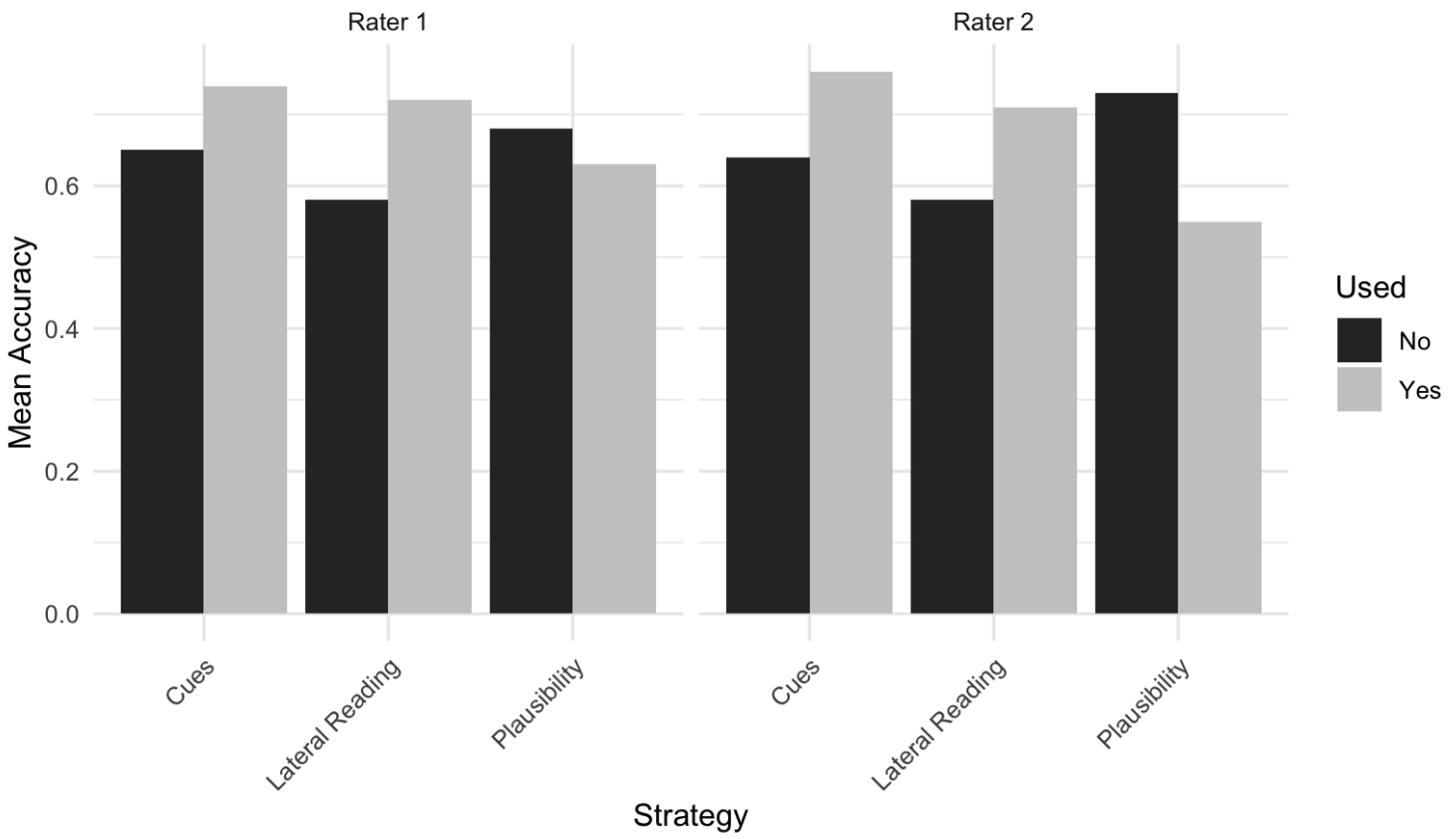


Figure 2

Mean accuracy by evaluative strategy. The figure compares the mean accuracy of student judgments based on whether each evaluation strategy (cues, lateral reading, plausibility) was used. Lateral reading is associated with higher accuracy for both raters, while plausibility-based reasoning corresponds to lower accuracy, especially for Rater 2. Cue-based reasoning shows moderate accuracy and less variation. These trends suggest that lateral reading is the most reliable strategy, whereas plausibility heuristics may lead to more frequent misjudgments.

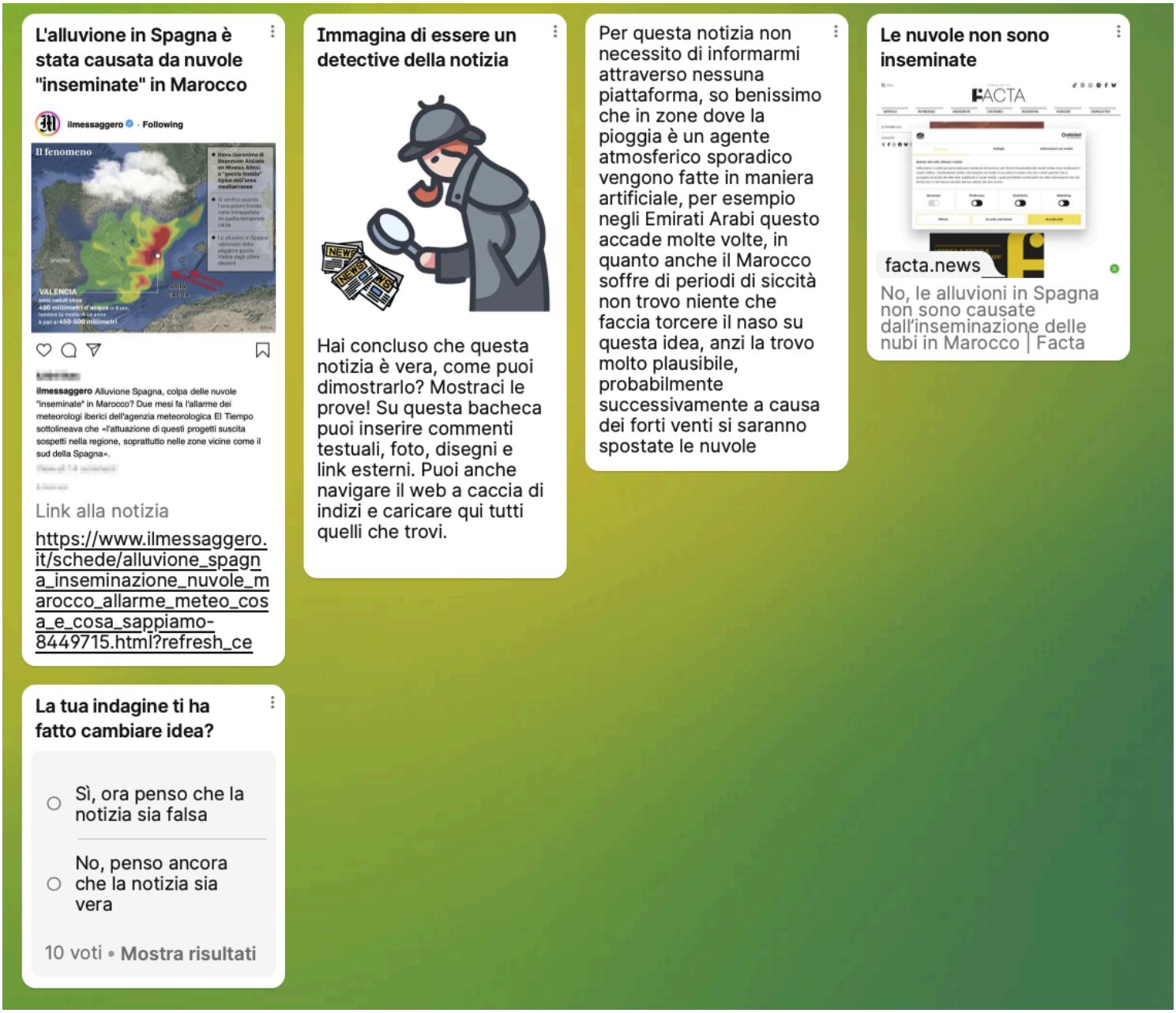


Figure 3

A snapshot of the Padlet board. An example of a Padlet board filled with a few comments shared by students.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryinformation.pdf