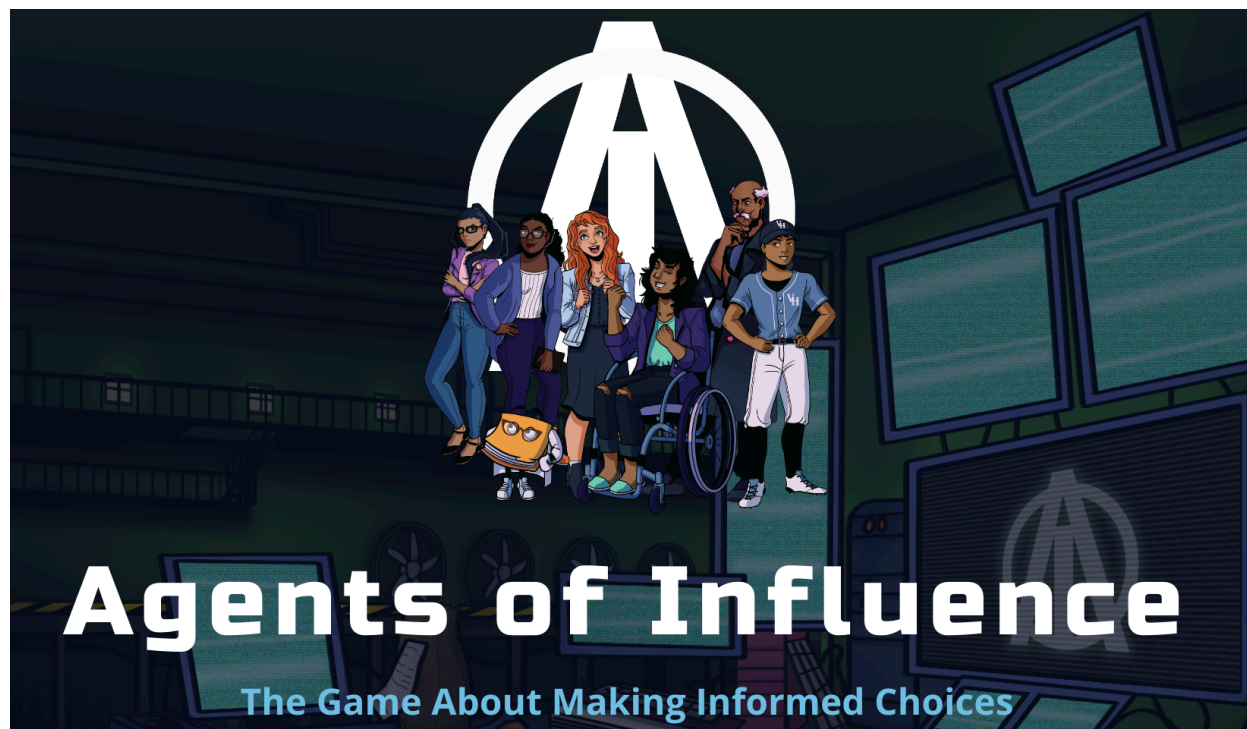


Agents of Influence

Improving Youth Resilience to Misinformation through a Gamified Digital Media Literacy Intervention



Altearea Inc. & Stanford Social Media Lab
White Paper

Executive Summary. Agents of Influence is an online video game designed to teach middle and high school students digital media literacy knowledge and critical thinking skills to improve their ability to recognize online misinformation and deception. Results from a longitudinal quasi-experimental study revealed that playing Agents of Influence improved diverse elementary school students' ability to sort fact from fiction by 30%. Students who played the game increased their knowledge of key skills like fact-checking, reverse-image searches, and lateral reading. They also improved their ability to critically evaluate information from websites promoting false information about climate change. Students enjoyed the game and felt more capable of identifying truthful, reliable information. Together, these findings indicate that Agents of Influence is an effective tool for building youth resilience to misinformation.

Overview:

The misinformation crisis. The internet and social media platforms have become rife with misinformation and disinformation. Concerns about misinformation exacerbate many of the key problems facing America today: intensifying polarization, promoting vaccine hesitancy, and undermining faith in democratic institutions. This has been amplified by the affordances of technology platforms, which allow falsehoods to proliferate – and mutate - at an unprecedented scale.

The Intervention: Agents of Influence, a media literacy game to teach middle and high school students how to find reliable information online and make responsible, informed decisions in an AI and social-media dominated world. The game is flexibly designed and can be used in various contexts including CTE, Expanded Learning, Libraries and in classrooms (it aligns with classroom standards in Technology, English, Social and Science.) The key skills it teaches - critical thinking, content analysis, research and communication - are the bedrock of education and necessary for thriving in the 21st Century.

The Results: Overall, the game was effective in enhancing students' discriminant trust—recognizing both true and false news- with their accuracy rising from 36% to 65% after playing the first episode. However, this improvement decreased over time, with only marginal gains after the second episode, leading to a 44% accuracy rate. Students who played the game increased their knowledge of key skills like fact-checking, reverse-image searches, and lateral reading. They also improved their ability to critically evaluate information from problematic websites. Students enjoyed the game and felt more capable of identifying truthful, reliable information.

Background.

The misinformation crisis. The internet and social media platforms have become rife with misinformation and disinformation. Concerns about misinformation exacerbate many of the key problems facing America today: intensifying polarization, promoting vaccine hesitancy, and undermining faith in democratic institutions. This has been amplified by the affordances of technology platforms, which allow falsehoods to proliferate – and mutate - at an unprecedented scale.

Skills-based interventions. While technological and regulatory progress are all required to combat misinformation, we also need a major investment in educating our citizenry to become resilient to misinformation. Nearly half of all Americans were exposed to at least one disinformation website in the lead-up to the 2016 election - amounting to around 150 million people (Guess et al., 2020). In their 2021 report (Aspen Digital, 2021), the Aspen Institute Commission on Information Disorder finds misinformation functions as a “force multiplier” exacerbating the worst problems of society nationally and globally, estimating that “hundreds of millions of people pay the price, every single day, for a world disordered by lies.”

Supporting American youth. Children and adolescents are also seriously impacted by misinformation. This presents a challenge for the next generation of young people who must learn to navigate an information ecosystem polluted with false and misleading information before they have come of age.

Young people are uniquely susceptible to believing and sharing misinformation. They tend to “info-graze” - meaning that they incidentally absorb news-related content and claims in the context of their existing social media use, as opposed to seeking out credible news-related information to answer relevant inquiries or fact-check claims (McGrew et al., 2018). Their reliance on social media as a primary source of news thus makes them susceptible to picking up and internalizing false claims in an uncritical manner (Oeldorf-Hirsch, 2018; Common Sense Media, 2023). Indeed, the first nationally representative study of young Americans’ recognition of misinformation paints a grim picture: even with access to the Internet, two-thirds of adolescents were unable to distinguish factual news articles from deceptive content because they were fooled by surface-level signals of credibility (Breakstone et al., 2021). Furthermore, 91% of AP history-enrolled high school students in a study trusted the claims of a highly partisan news site (Weinberg et al., 2022), while surveyed teens listed entertainers and YouTubers among their most trusted news sources (Common Sense Media, 2023). It is clear we need to support youth in learning how to navigate information online.

Skills-based approaches to building misinformation resilience. Early research suggests that relatively brief digital media literacy interventions have the potential to dramatically improve people’s ability to discern credible information online from misinformation (Guess et al., 2020; Walther et al., 2022; Lee et al., 2024). These interventions work by equipping individuals with the skills they need to discern between true and false

information online, such as by employing lateral reading strategies to evaluate the quality of claims across multiple and using reverse image searches. The Aspen Institute argues that digital media literacy interventions are essential to our national disinformation response: “Users need to have the tools that can help them distinguish fact from falsehood, honesty from manipulation, and the trustworthy from the fringe” (p.66).

The potential for gamified, interactive media experiences. While existing digital media literacy interventions show promise in protecting individuals from the harms of misinformation, more work is needed to rigorously evaluate their impact on beliefs and behavior among youth and adolescents. The vast majority of interventions to date have been designed for adult populations, often in the format of multi-hour Zoom workshops or self-paced online courses. Students - particularly children and adolescents - may struggle to stay engaged with such content and may instead benefit from interventions tailored to be developmentally appropriate and engaging.

Gamification is a growing tool in counter-misinformation interventions, as it provides educational benefits while appealing to the new generation’s digital engagement patterns. Some of the research-based benefits of gamification include high transferability of skills and increased skill retention through emotional resonance with characters and storylines (Pappas, 2015). Developing a gamified digital media literacy interventions can make a significant impact on teenage disinformation resilience by being fun (enjoyability, memorability, connection to students); educational (covers meaningful information, does so effectively, easy to use in the classroom); and robust (produces change in students’ ability to detect misinformation, trust credible information, and take action against false claims).

About the Game.

Game Overview

The story and name of *Agents of Influence* come from “agent of influence,” a Cold War-era term describing any spy who influenced foreign policy with misinformation. The game reclaims the term by showing students have “agency” (ability to make independent decisions) and “influence” (impact on others through their decisions), both of which are essential to be informed and empowered citizens.

Story

The game is set at Virginia Hall high, a fictional high school overrun with misinformation schemes. In this narrative adventure, you, your friends, and your grandpa Teddy Turn, an ex-secret agent, start a spy organization called the Agents of Influence to save the school from Teddy's archnemesis Nepravda and his shadowy organization, Harbinger.

Game Mechanics

Digital skills are taught and reinforced through a series of mini-games:

1. Conversation Game: Players learn how to engage in conversations with antagonistic individuals by avoiding logical fallacies, treating others with respect, and selecting dialogue choices that move the conversation forward.
2. Research Game: This game teaches players how to ask good questions and investigate claims across platforms to verify their trustworthiness.
3. Document Analysis Game: Players critically evaluate media by answering true or false questions to determine the content's credibility.

By playing these games, students will develop and strengthen their abilities to protect themselves and their communities from disinformation.



Figure 1. In the Research Game, students are sent on a field mission to investigate information online. They must work together with their AI Assistant, A.M.I.E., and apply their media literacy skills to determine which sources are relevant and reliable, and which are not.

The Research Study.

The Stanford research team conducted a quasi-experimental study to examine how playing the Agents of Influence video game influenced students' ability to recognize and respond to false, misleading, or deceptive information they may encounter online. Students were asked to take three surveys assessing their media literacy skills and ability to detect misinformation online (1) before playing the game, (2) after playing Episode 1, and (3) after playing Episode 2.

Ninety-three elementary school students from public, private, and home-schools in California to play Agents of Influence. There were 50 boys (53.8%), 42 girls (45.2%), five students that were non-binary or questioning, and one student that declined to indicate their gender. The sample was predominantly white ($n = 68$, 73.1%), with 14 students who identified as Asian/Asian American (15%), one student that was Hispanic/Latino (1%), and four students that preferred not to disclose their ethnicity. The average age was 13 years old, with students ranging from 11 to 14.

Key results.

Students improved core digital media literacy skills. Playing the Agents of Influence game significantly boosted students' digital media literacy over time. Results from a series of linear mixed effects regression models controlling for demographic differences in student age, gender, and school revealed that students improved their understanding of how to use fact-checking resources to evaluate claims, how to conduct reverse image searches to understand where photos come from, and how to identify sources of trustworthy, reliable, and relevant information. After playing the Research Game, students also specifically improved their knowledge of good questions to ask to interrogate online claims and their ability to evaluate the credibility of online news sources. There were no changes in skills like lateral reading or understanding social media algorithms. There were slight decreases in students' reported knowledge of how to use search engines over time.

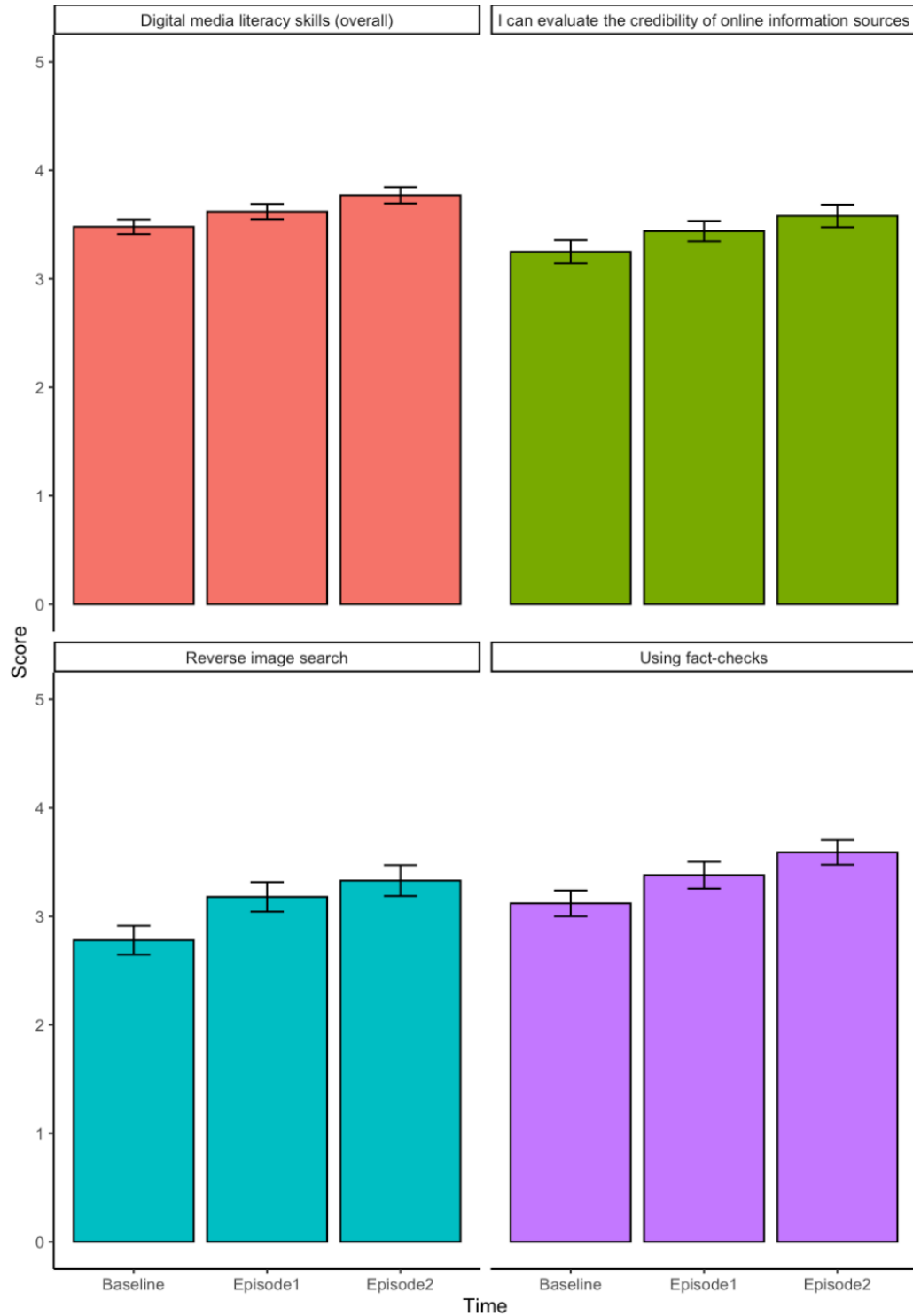


Figure 2. Changes in digital media literacy skills over time. Red = overall digital media literacy skills composited over six variables (knowledge of lateral reading, click restraint, monitoring emotional reactions to headlines, reverse image search, and using fact-checking resources, finding credible information online), green = ability to evaluate

credibility of online information, blue = reverse image search, purple = using fact-checks.

Students are better at detecting misinformation online. The study found that students' ability to accurately discern trustworthy from untrustworthy news improved significantly after playing the first episode of the game. We asked them to evaluate the credibility of some headlines which had been fact-checked to be true or false and analyzed how many they got right. Overall, the game was effective in enhancing students' discriminant trust—recognizing both true and false news- with their accuracy rising from 36% to 65% after playing the first episode. However, this improvement decreased over time, with only marginal gains after the second episode, leading to a 44% accuracy rate. There were no significant differences in outcomes based on students' age, ethnicity, or gender.

One reason they may have improved their ability to detect misinformation over time is because they were also applying the digital media literacy skills they had learned. For each headline that they were asked to evaluate, they were also asked if they had done any research to investigate the veracity of the claim. While students only investigated 23% of headlines at baseline on average, they were more likely to use their digital media literacy skills after playing the first episode of Agents of Influence - investigating up to 34% of the headlines they saw.

The game also improved students' ability to evaluate the veracity of online sources. Following McGrew et al. (2022), students were asked to investigate the credibility of several real websites that are known for producing climate change misinformation. This is an important measure given that students often need to evaluate information in the context of their online web-browsing activity. Results from the first model indicated that students' improved their ability to identify low-credibility websites after playing episode 2, but not episode 1. This pattern of results indicated that the Research Game significantly improved students' ability to not be fooled by these websites.

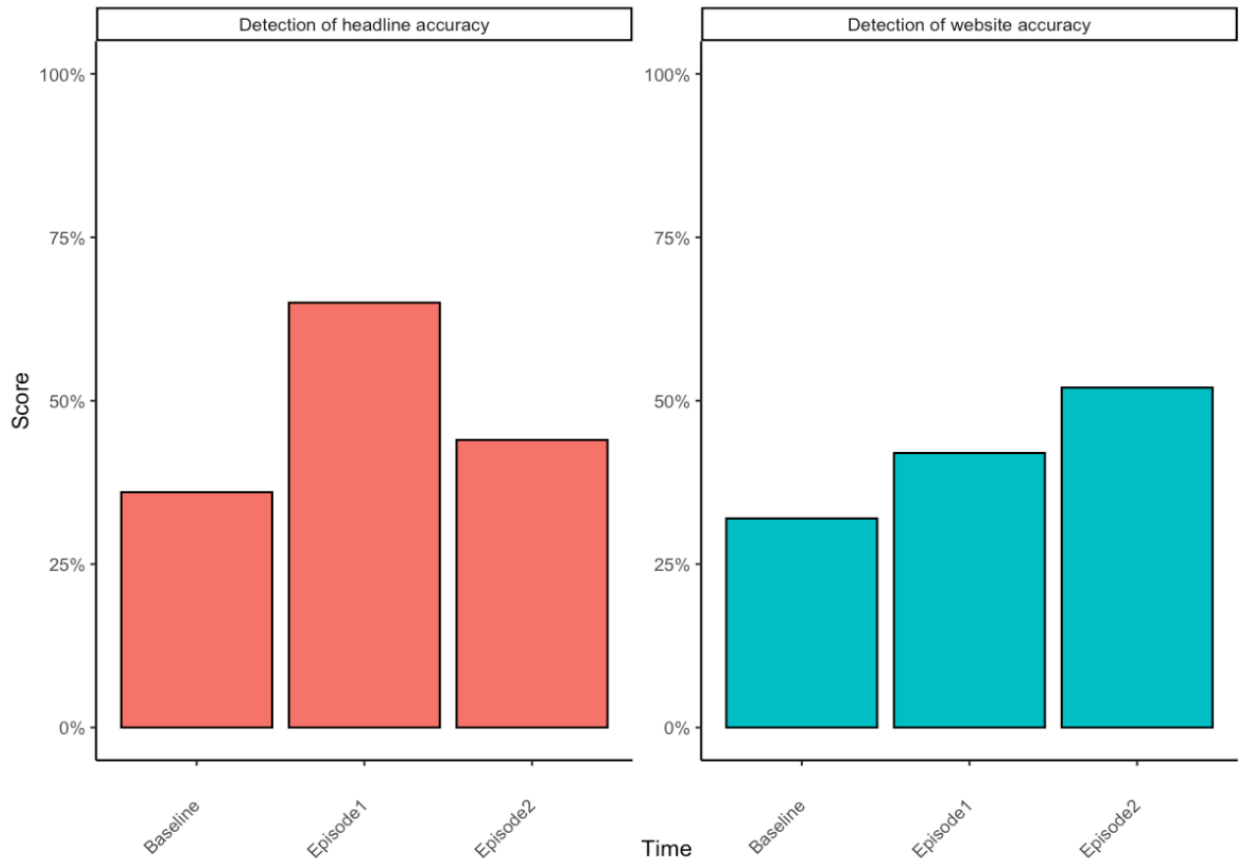


Figure 3. Changes in students' ability to detect misinformation and website accuracy over time.

Students feel more confident in their own skills and enjoyed the game. Crucially, playing Agents of Influence increased students' sense of self-efficacy in tackling misinformation and deception online. Relative to baseline, students felt more confident in their own abilities after playing Agents of Influence. Students also reported having fun playing the game, saying they learned from it, enjoyed the experience, and felt more capable of identifying truthful, reliable, and relevant information.

Implications and Significance.

This study is crucial because it addresses the urgent need to equip the next generation with effective tools to navigate the complex and often deceptive online information environment. With misinformation increasingly influencing public opinion and societal issues such as polarization and vaccine hesitancy, teaching children and adolescents how to critically evaluate information is more important than ever. Existing research highlights that young people, despite their digital fluency, are particularly vulnerable to misinformation due to their tendencies to "info-graze" and the reliance on social media for news. By focusing on a novel, gamified intervention designed specifically for youth, this study aims to bridge the gap between traditional media literacy approaches and the engagement needs of younger audiences, thereby fostering greater resilience against digital misinformation.

Furthermore, the study's emphasis on a gamified digital media literacy intervention represents a significant advancement in combating misinformation among children and adolescents. Unlike conventional workshops and courses, the gamified approach is designed to be engaging and developmentally appropriate, potentially enhancing the retention and application of digital media literacy skills. By incorporating interactive and immersive elements, such as mini-games within a spy-themed narrative, the intervention seeks to not only educate but also captivate students, making the learning process both enjoyable and memorable. This innovative approach holds the promise of significantly improving students' ability to discern credible information from falsehoods, thus contributing to their long-term ability to navigate and mitigate the impacts of misinformation.

Notably, our results demonstrate that Agents of Influence does not only improve students' ability to correctly detect *false* information, but also improves their trust in *true news*. This characteristic is particularly important as solutions to the misinformation crisis should not come at the cost of reducing trust in *all* news or increasing skepticism across the board. Together, our findings attest to the potential for gamified misinformation interventions to build resilience against misinformation among young people.

Limitations and Future Research

The results of this study should be interpreted with respect to its limitations. The quasi-experimental nature of our study limits the internal validity of our study because the

lack of random assignment means that there may be characteristics of our recruited sample that influenced how they responded to the game. In addition, our method of recruitment means that there may be response bias in the students who chose to participate. Future research will conduct a randomized controlled trial with a nationally representative sample to experimentally test the impact of the Agents of Influence game on a larger, more diverse sample, with less information about the purpose of the study.

Table 1. Means of student skills and experiences before and after playing Agents of Influence

Table 2. Means before and after playing AOI

	Baseline	After Playing Episode 1	After Playing Episode 2
Digital media literacy skills (overall)	3.48 (.60)	3.62 (.63)	3.77 (.67)
Using fact-checks	3.12 (1.07)	3.38 (1.10)	3.59 (1.02)
Reverse image search	2.78 (1.19)	3.18 (1.22)	3.33 (1.27)
Using search engines	4.54 (.63)	4.38 (.79)	4.37 (.77)
Lateral reading	3.52 (.93)	3.62 (.96)	3.74 (.93)
Understanding algorithmic curation	3.27 (.96)	3.48 (.81)	3.64 (.84)
I can evaluate the credibility of online information sources	3.25 (.96)	3.44 (.84)	3.58 (.93)
I can evaluate the purpose of online information sources	3.53 (.88)	3.68 (.86)	3.64 (.84)
I can evaluate the main idea of online information sources	3.71 (.86)	3.74 (.86)	3.77 (.89)
I can evaluate when online information	3.57 (.91)	3.72 (.91)	3.72 (.93)

sources use evidence-based arguments			
<i>Discernment of true from false news</i>			
Identifying the markers of accurate information	3.26 (.93)	3.49 (.80)	3.70 (.85)
Detection of headline accuracy	36%	65%	44%
Detection of website accuracy	32%	42%	52%
<i>Enjoyment of game</i>			
How much did you enjoy the AOI game?	-	3.61 (.96)	3.79 (1.02)
How much did you learn from the AOI game?	-	2.92 (1.01)	3.29 (.93)
How capable do you feel of identifying truthful, reliable, and relevant information?	-	3.57 (.82)	3.79 (.82)
How much do you think that misinformation is an important problem?	-	3.97 (.97)	4.05 (.95)

Note. All variables are scored on a 5-point Likert scale. Bold items demonstrated significant improvement over time. Numbers in parentheses indicate the standard deviation.

Detailed Methodology.

Digital media literacy skills.

We assessed participants' digital literacy skills by asking participants to complete a 6-item measure rating their understanding of lateral reading, click restraint, monitoring emotional reactions to headlines, reverse image search, and using fact-checking resources. Following prior work, participants rated their comprehension of each skill on a 5-point scale, ranging from 1:no understanding - 5:full understanding (Lee et al., 2024; Hargittai & Hsieh, 2011).

To test if playing the game would improve students' digital media literacy skills over time, we fit a linear mixed effects regression model where the predictor variable was time (T1 = before game, T2 = after playing episode 1, T3 = after playing episode 2) and the dependent variable was the average of participants' digital media literacy skills. We included student gender, age, and ethnicity as covariates to control for potential demographic differences. A random intercept was included for participants to account for each student completing multiple assessments over time.

Relative to baseline, students' comprehension of digital media literacy skills improved after playing the first episode of AOI ($\beta = .13$, $SE = .06$, $p = .03$) and after playing the second episode of AOI ($\beta = .25$, $SE = .06$, $p < .001$). A post-hoc comparison revealed that the improvement in students' skills from T2 to T3 was significant ($\beta = .13$, $SE = .05$, $p = .009$). Older students ($\beta = .17$, $SE = .06$, $p = .01$) and non-white students ($\beta = .35$, $SE = .15$, $p = .02$) demonstrated greater gains in digital media literacy skills.

A closer examination revealed that this effect was driven by changes in specific skills. We fit a series of seven, separate models following the same specifications where the dependent variable was each digital media literacy skill. Results indicated that the intervention was effective at improving students' knowledge of fact-checking resources (after episode 1: $\beta = .24$, $SE = .12$, $p = .053$; after episode 2: $\beta = .49$, $SE = .13$, $p < .001$); reverse-image searches (after episode 1: $\beta = .36$, $SE = .13$, $p = .005$; after episode 2: $\beta = .45$, $SE = .13$, $p < .001$); and the markers of trustworthy information online (after episode 1: $\beta = .25$, $SE = .10$, $p = .012$; after episode 2: $\beta = .37$, $SE = .10$, $p < .001$). Students' knowledge of the questions they could ask to consider claims improved after playing episode 2 ($\beta = .37$, $SE = .11$, $p < .001$), but not after playing episode 1 ($\beta = .17$, $SE = .10$, $p = .09$). No significant changes were observed over time for lateral reading (after episode 1: $\beta = .10$, $SE = .12$, $p = .41$; after episode 2: $\beta = .16$, $SE = .12$, $p = .17$) or for how social media algorithms affect what people see online (after episode 1: $\beta = -.03$, $SE = .10$, $p = .72$; after episode 2: $\beta = .12$, $SE = .10$, $p = .24$). Knowledge of how to use search engines decreased significantly over time (after episode 1: $\beta = -.19$, $SE = .09$, $p = .04$; after episode 2: $\beta = -.19$, $SE = .09$, $p = .04$).

Additional skills. The AOI game also increased students' ability to evaluate the credibility of online information sources after Episode 2 ($\beta = .29$, $SE = .11$, $p = .003$). It did not improve students' ability to evaluate the purpose of online information sources, the main idea of online information sources, or their ability to evaluate when online sources used evidence-based arguments.

Discriminant trust and ability to recognize misinformation.

As in most studies of misinformation interventions, we evaluated participants' ability to identify true and false claims online by asking participants to judge the veracity of a series of headlines (Pennycook et al., 2021). At each time point, students saw two news headlines: one of which was false and one of which was true. All headlines were determined to be either true or false by reputable fact-checking organizations such as AP News, Reuters, or Snopes. For each headline, participants reported how true or false they thought each headline was (scale from 1: definitely false to 7: definitely true). The accuracy of each judgment was calculated by coding headline evaluations as correct if individuals rated false news as definitely to somewhat false (1-3 on a 7-point scale), and true news as definitely to somewhat true (5-7 on a 7-point scale); responses of 4 were coded as incorrect. Afterwards, participants were provided with a fact-sheet informing them which headlines they viewed were true and false, with links to fact-checking articles. To test if students applied their digital media literacy skills, they were asked after submitting their headline judgments if they had done any research on the claim before making their evaluation (1 = yes, 0 = no).

A mixed effects logistic regression model was estimated to evaluate whether playing the Agents of Influence game (1 = before game, 2 = after episode 1; 3 = after episode 3) influenced students' ability to accurately identify true and false headlines (1 = judgment of a headline was accurate, 0 = judgment of a headline was inaccurate). Random effects were modeled to account for each participant reporting multiple occasions of data, and each participant six headlines in total.

Results indicated that students' discriminant trust improved significantly after playing the first episode of the game ($\beta = 1.25$, $SE = .23$, $p < .001$) from 36% accuracy ($SD = .34$) before the intervention to 65% ($SD = .32$). However, performance diminished over time such that there were only marginally significant improvements from baseline after playing episode two ($\beta = .43$, $SE = .23$, $p = .06$) (44% overall accuracy). There were no significant differences for students of different ages, ethnicities, or genders. Students' detection of true headlines increased from 29% at baseline to 66% after the first episode ($\beta = 2.25$, $SE = .45$, $p < .001$) and 43% after the second episode ($\beta = .88$, $SE = .40$, $p = .03$). Similarly, students' ability to correctly identify false news as false increased from 44% at baseline to 65% after playing the first part of AOI ($\beta = .85$, $SE = .35$, $p = .01$). However, they returned to 45% ($\beta = .21$, $SE = .34$, $p = .54$). Results indicated that the intervention was effective at improving students' *discriminant trust* - their ability to both recognize true news as true, and false news as false - over time, but their ability to retain trust in true news was sustained for longer.

We fit another mixed effects linear regression model where the dependent variable was the students' average rate of doing research on the claim, at each time point. The same covariates and random effects were included as above. While students only investigated 23% of headlines at baseline ($SD = 37\%$) on average, they were more likely to use their digital media literacy skills after playing the first episode of Agents of Influence - investigating up to 34% of the headlines they saw ($SD = 44\%$) ($\beta = .10$, $SE = .04$, $p = .01$). However, this effect was no longer significant after time two with students only investigating 29% of headlines ($\beta = .07$, $SE = .04$, $p = .09$).

Finally, we tested if the intervention was effective at teaching students to critically evaluate the veracity of online sources. Following McGrew et al. (2022), students were asked to investigate the credibility of several websites that are known for producing climate change misinformation. We fit a pair of mixed effects logistic regression with the same specifications as above, where the dependent variables were 1) students' ability to correctly identify if the website promoted false information (1 = correct) or not (0 = false), and 2) their *mastery* of the digital media literacy skills they were taught (see McGrew et al., 2018 and McGrew et al., 2022). Results from the first model indicated that students' improved their ability to identify low-credibility websites after playing episode 2 of AOI ($\beta = 1.26$, $SE = .41$, $p = .002$), but not after episode 1 ($\beta = .56$, $SE = .38$, $p = .14$).

Discriminant trust and ability to recognize misinformation.

We fit a pair of mixed effects regression models where the predictor variables were time and the dependent variables were 1) a subjective assessment of their own ability to recognize misinformation online, and 2) the mean of six items that measured their self-efficacy regarding misinformation detection, respectively based off of the New General Self-Efficacy Scale (Chen et al., 2001). The same covariates and random effects were included as above.

Results indicated that students became more confident in their own abilities to recognize misinformation after finishing the intervention (after episode 1: $\beta = .22$, $SE = .06$, $p < .001$; after episode 2: $\beta = .24$, $SE = .06$, $p < .001$). Girls also indicated they were more confident than boys ($\beta = .33$, $SE = .14$, $p = .02$). However, no significant changes in students' self-efficacy were observed over time (after episode 1: $\beta = .09$, $SE = .08$, $p = .26$; after episode 2: $\beta = .06$, $SE = .08$, $p = .45$).

Conflict of Interest Statement.

Data analysis was conducted independently by AL of the Stanford Social Media Lab. There were no financial conflicts of interest to report.

References.

- Breakstone, J., Smith, M., Wineburg, S., Rapaport, A., Carle, J., Garland, M., & Saavedra, A. (2021). Students' civic online reasoning: A national portrait. *Educational researcher*, 50(8), 505-515.
- Couric, K., Krebs, C., & Robinson, R. (2021). Commission on Information Disorder Final Report. November, Washington, DC: Aspen Institute.
- Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors

of fake news dissemination on Facebook. *Science advances*, 5(1), eaau4586.

Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences*, 117(27), 15536-15545.

Lee, A. Y., Moore, R. C., & Hancock, J. T. (2024). Building resilience to misinformation in communities of color: Results from two studies of tailored digital media literacy interventions. *new media & society*, 14614448241227841.

McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & research in social education*, 46(2), 165-193.

Oeldorf-Hirsch, A. (2018). The role of engagement in learning from active and incidental news exposure on social media. *Mass communication and society*, 21(2), 225-247.

Pappas, G., Papamichael, I., Zorpas, A., Siegel, J. E., Rutkowski, J., & Politopoulos, K. (2021). Modelling key performance indicators in a gamified waste management tool. *Modelling*, 3(1), 27-53.

Walther, J. B., Lew, Z., Edwards, A. L., & Quick, J. (2022). The effect of social approval on perceptions following social media message sharing applied to fake news. *Journal of Communication*, 72(6), 661-674.

Wineburg, S., Breakstone, J., McGrew, S., Smith, M. D., & Ortega, T. (2022). Lateral reading on the open Internet: A district-wide field study in high school government classes. *Journal of Educational Psychology*, 114(5), 893.