Fakey: A Game Intervention to Improve News Literacy on Social Media

NICHOLAS MICALLEF, Center for Cybersecurity, New York University Abu Dhabi, United Arab Emirates
MIHAI AVRAM, Observatory on Social Media, Indiana University Bloomington, USA and Luddy School of Informatics, Computing, and Engineering, Indiana University Bloomington, USA
FILIPPO MENCZER, Observatory on Social Media, Indiana University, USA and Luddy School of Informatics, Computing, and Engineering, Indiana University Bloomington, USA
SAMEER PATIL, Luddy School of Informatics, Computing, and Engineering, Indiana University Bloomington, USA

We designed and developed Fakey, a game to improve news literacy and reduce misinformation spread by emulating a social media feed. We analyzed player interactions with articles in the feed collected over 19 months within a real-world deployment of the game. We found that Fakey is effective in priming players to be suspicious of articles from questionable sources. Players who interact with more articles in the game enhance their skills in spotting mainstream content, thus confirming the utility of Fakey for improving news literacy. Semi-structured interviews with those who played the game revealed that players find it simple, fun, and educational. The principles and mechanisms used by Fakey can inform the design of social media functionality to help people distinguish between credible and questionable content in their news feeds.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**; *Empirical studies in collaborative and social computing*.

Additional Key Words and Phrases: misinformation, fake news, low-credibility content, news literacy, game, social media, news feed

ACM Reference Format:
Nicholas Micallef, Mihai Avram, Filippo Menczer, and Sameer Patil. 2021. Fakey: A Game Intervention to Improve News Literacy on Social Media. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 6 (April 2021), 27 pages. https://doi.org/10.1145/3449080

1 INTRODUCTION

As a large proportion of the population receives news from social media, online misinformation has emerged as a critical societal threat. Nearly 66% of American adults stated that "fake news" has caused them confusion, and nearly a quarter admitted to sharing such content, knowingly or unknowingly.[1] The spread of misinformation can have severe consequences. For instance, several

[1] https://www.pewresearch.org/fact-tank/2017/12/28/key-trends-shaping-technology-in-2017/

6

Authors' addresses: Nicholas Micallef, Center for Cybersecurity, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, nicholas.micallef@nyu.edu; Mihai Avram, Observatory on Social Media, Indiana University Bloomington, Bloomington, IN, USA, Luddy School of Informatics, Computing, and Engineering, Indiana University Bloomington, Bloomington, IN, USA, mihai.v.avram@gmail.com; Filippo Menczer, Observatory on Social Media, Indiana University, Bloomington, IN, USA, Luddy School of Informatics, Computing, and Engineering, Indiana University Bloomington, Bloomington, IN, USA, fil@indiana.edu; Sameer Patil, Luddy School of Informatics, Computing, and Engineering, Indiana University Bloomington, Bloomington, IN, USA, patil@indiana.edu.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs International 4.0 License.

© 2021 Copyright held by the owner/author(s).
2573-0142/2021/4-ART6. https://doi.org/10.1145/3449080

Telecom engineers in the United Kingdom received death threats from people who believed that the novel coronavirus was being spread through newly installed 5G antennas,[2] and 5G antennas were set on fire in several locations.[3] Recently, election-related misinformation was a major factor underlying the deadly attack on the US Capitol.[4]

The impact of misinformation could be substantially reduced if people readily recognized and ignored such content. Researchers have attempted to address this issue by promoting greater news literacy via various means, such as teaching people to recognize fake news articles [5, 64], changing people's news consumption behavior [14, 38, 42, 45], and flagging questionable content with credibility indicators [53, 57, 58, 84].

The prevalence of the spread of misinformation within social media news feeds has been examined to determine the impact of various relevant signals, such as news sources [46], social engagement metrics [4], text and image cues [29, 67], etc. For instance, the literature provides consistent evidence that social media users do not pay much attention to the source of the news [13, 33, 45, 52]. While there have been explorations for enhancing the news feed with new signals [6, 31], there has been little empirical research on training social media users to be vigilant toward misinformation and to use available signals to recognize and scrutinize suspicious content [58].

To address the research gaps identified above, we designed, developed, and deployed *Fakey*,[5] a game that emulates relevant interface elements of popular social media platforms such as Facebook and Twitter. Fakey simulates a news feed in which players can `Share`, `Like`, or `Fact Check` individual articles (see Figure 1). Each article in the news feed contains a headline, a photo, a description, and a *simulated* social engagement metric randomly generated with a distribution similar to that for engagement metrics encountered on real-world social media services. The articles are randomly selected in equal measure from mainstream (e.g., The New York Times) and low-credibility (e.g., Now8News) sources. The game was designed to answer the following research question: *Can a game that emulates social media feeds improve news literacy?*

We addressed the above question by analyzing the analytics of user interactions collected over 19 months in a real-world deployment of the game. We separately conducted a small-scale investigation where we interviewed people immediately after they played the game. We found that Fakey helps players improve news literacy. As they play the game, players get better at recognizing mainstream content. The interviews confirmed that players understand the purpose of the game and find it useful.

Specifically, we make the following contributions:

- We present a game to promote news literacy and deploy it in the real-world.
- We demonstrate that players who interact with more articles in the game improve their capability to recognize mainstream content without affecting the ability to spot low-credibility content.

In the sections that follow, we begin by situating our work in the literature regarding news literacy, misinformation, social news feed signals and related gaming, behavioral, and corrective interventions. We then describe the operation of Fakey along with our data collection and analysis approaches. Next, we report the findings related to our research question, followed by a discussion of the practical implications of the insight. At the end, we mention a few limitations and conclude with important future directions.

[2]https://www.bbc.com/news/newsbeat-52395771
[3]https://www.bbc.com/news/uk-england-leeds-52692654
[4]https://apnews.com/article/donald-trump-conspiracy-theories-michael-pence-media-social-media-daba3f5dd16a431abc627a5cfc922b87
[5]https://fakey.iuni.iu.edu/

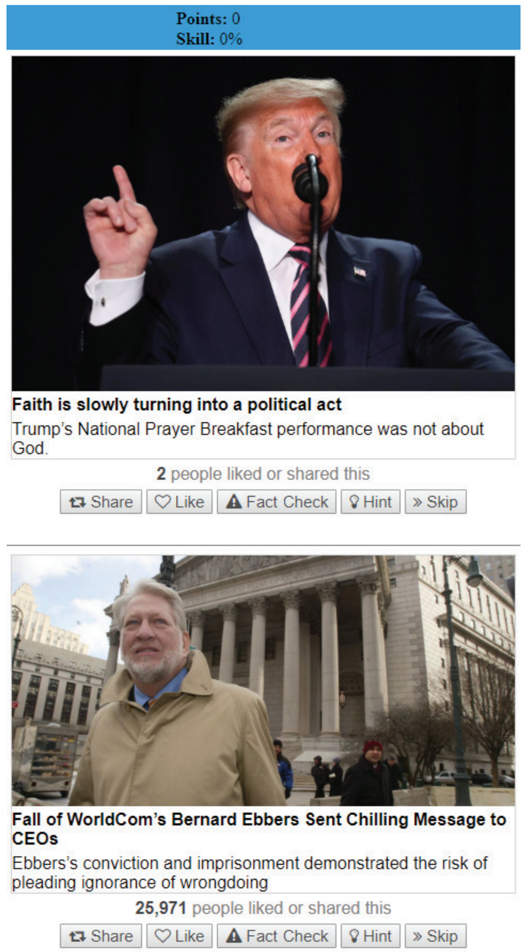


Fig. 1. Example news feed as seen within Fakey.

2 RELATED WORK

News literacy is defined as the ability to access, analyze, evaluate, and create messages in a variety of forms [3]. This definition has been extended to include the heterogeneous and dynamic environment of online media [41]. The exploding volume of online news makes it challenging to engage in rigorous fact checking. As a consequence, there has been a rise in misinformation and propaganda aimed at influencing or deceiving people [32, 38]. In the following subsections, we cover the salient existing work on the prevalence of online misinformation embedded within social media platforms along with interventions to shape user behavior to detect and curb its spread.

2.1 Misinformation Prevalence and Spread

Research on misinformation includes methods to quantify its prevalence [72, 75]. Prevalence has been measured for various types of misinformation, such as false information [2, 28, 71], rumors [18, 70, 73], deepfake videos [10, 23], etc. Although a majority of misinformation remains obscure, the small percentage of it that spreads does so more virally than true information [79].

The viral spread of misinformation has increased the pressure on journalists and fact checkers who typically lack the resources to verify all of the false claims spread online [50].

Misinformation would be less prevalent if information consumers could easily recognize and ignore it. In reality, people struggle to distinguish between true and false information online [36, 37, 61], discerning misinformation with an accuracy ranging from 53% to 78% [37, 61]. Older users (ages 65 or above) share seven times as many false news articles on average compared to younger users [25]. Long, well written, and craftily referenced misinformation can fool even those who have been trained to spot false claims [36]. Social media further exacerbates the vulnerability to misinformation because users of these platforms are affected by groupthink, filter bubbles, and echo chambers [12, 15, 20, 62]. Other factors that lower the ability to recognize misinformation include low levels of education, low media consumption, confirmation bias, and conservative political beliefs [24, 51, 54, 60, 69].

2.2 Social News Feed Signals

To reduce the spread of misinformation via social media, researchers have examined news consumption practices on these platforms, paying particular attention to the various signals present in the User Interface (UI) of the news feeds. For instance, Glenski et al. [22] found that most Reddit users read only the headlines and the corresponding summary, without accessing the full content or reading the associated comments. Although some researchers have found that social media users utilize the available cues pertaining to the source of a news story [13, 74], most studies report that users do not pay much attention to source information [13, 33, 45, 52]. Indeed, those who come across a news story on social media are far less likely to attribute it to the original source than those who encounter the same story directly on the site of the source [34].

Rather than basing their evaluation of a news article on its original source, social media users are instead more likely to determine trustworthiness of the story depending on whether it is shared or endorsed by a well-known public figure and, in turn, more likely to endorse and/or share it with their friends or family [74]. Since people place a great deal of importance on such social endorsements when selecting the content to consume on social media, they are much less likely to be exposed to content that covers diverse perspectives, such as viewpoints or arguments that counter misinformation [46]. Indeed, social engagement metrics, such as the number of Likes, have been found to increase vulnerability to misinformation [4]. Relatedly, the use of images alongside text can boost the credibility of posted content, regardless of veracity [29].

While the above research has examined the impact of the available signals on credibility judgments, the application of these signals to train users to recognize and scrutinize suspicious content [58] has received limited attention. Our research addresses this gap.

2.3 Behavioral and Corrective Interventions

Various behavioral interventions have been attempted to improve people's news literacy. These include: (i) encouraging people to pause to assess the credibility of headlines [6, 14]; (ii) promoting civic online reasoning [44] and critical thinking [42, 45]; (iii) enhancing social media feeds with signals derived from the posting history of the source [31]; and (iv) adding credibility indicators and warnings to prompt users to question or refute the information being consumed [7, 11, 53, 56–58, 80, 84].

In addition, researchers have evaluated the use of corrective interventions aimed at reducing the spread of misinformation. Such studies have found that succinct and repeated corrections can reduce misperceptions created by false information [40, 80], even for users who report disregarding the correction [7] and for misinformed users who might be affected by confirmation bias [76]. The corrections that have a measurable impact on social media user behavior are those that come from misinformation-flagging algorithms [7, 11, 57], reputable organizations (e.g., the CDC) [80, 82], domain experts [27], or multiple other users who link to a credible source [43, 81]. The outcomes of professional fact checking efforts can be similarly effective depending on the manner in which they are conveyed to users [1, 21, 54, 77, 83–85].

Although corrections can address misperceptions created by misinformation, they might not always change the underlying attitudes and beliefs [53, 77]. Another limitation of typical corrective mechanisms is that their implementation within real-world platforms does not necessarily ensure that they reach the users who were exposed to the corresponding misinformation [26, 30, 49, 68].

2.4 Game Interventions

Some research efforts have involved the use of games as interventions to improve news literacy. For instance, the game PolitiTruth[6] is similar to the dating app Tinder. PolitiTruth presents players with statements of public figures along with a timestamp. Players swipe left to classify the statement as fake or right to classify it as true. However, the impact of the game is limited to quotes from public figures and does not cover a broad spectrum of diverse misinformation content. In contrast, Factitious[7] is a game that provides players with the text of a short article, along with the corresponding image, headline, and source. After indicating whether they believe that the article is true, players are given the option to see the article in the context of its original source and an explanation of the correct choice. While useful as basic news literacy training, the limited set of manually curated articles available within the game significantly limits its scope. Play Fake News[8] uses content similar to Factitious and user interaction similar to PolitiTruth, thus suffering from the same limitations as those two games.

Instead of focusing on fake news *consumption*, the Fake It To Make It[9] and Bad News[10] games simulate the experience of fake news *creation*, to teach players about various misinformation strategies. For instance, players of Bad News learn six common techniques used to produce misinformation: creating polarization, invoking emotions, spreading conspiracy theories, trolling people, deflecting blame, and impersonating fake accounts [63, 64]. The premise of Fake It To Make It and Bad News is that learning about the tactics of misinformation creators can help people spot them in the content they consume. Given that misinformation spread shares commonalities with viral contagion [9], the premise is an attempt to apply inoculation theory to create a misinformation 'vaccine' that creates 'mental antibodies' by exposing people to existing misinformation practices [78]. Although there has been no reported evaluation of Fake It To Make It, a large-scale evaluation of Bad News showed that it helps people improve their ability to recognize fake news [5, 64]. However, both games currently include limited content that lacks variety. Therefore, the extent to which their premise is successful in a real-world environment is yet to be confirmed.

In contrast to the above games, Fakey covers a large volume of diverse content because it incorporates *live* news articles updated dynamically, thus ensuring that players always obtain previously unseen news articles. Moreover, Fakey is more directly connected to the real-world context because of its emulation of social media news feeds.

[6] https://www.cinqmarsmedia.com/politifact/
[7] http://factitious.augamestudio.com/
[8] https://www.rand.org/research/projects/truth-decay/fighting-disinformation/search/items/fake-news-the-game.html
[9] https://www.fakeittomakeitgame.com/
[10] https://getbadnews.com/

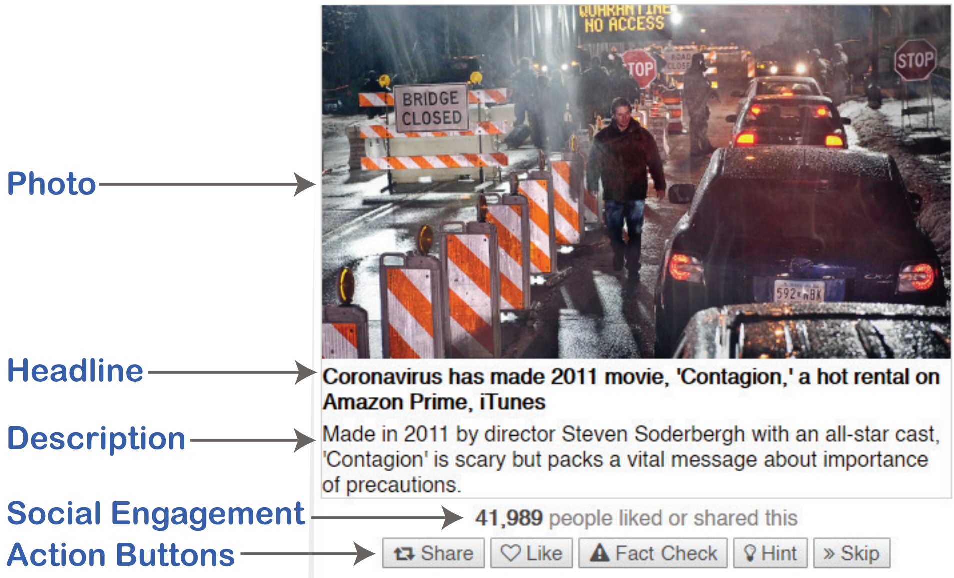


Fig. 2. An item in the news feed as presented within Fakey.

3 METHODS

Fakey was released to the general public as a web and mobile app. We analyzed the analytics collected by public use of Fakey followed by a small-scale study in which we conducted semi-structured interviews to verify player understanding of the game elements. The procedures for both components of the research were reviewed and approved by the Indiana University Institutional Review Board (IRB).

3.1 Game Design

Fakey emulates a social media news feed with the aim of analyzing user interaction with news items. The UI of the game mimics the appearance of Facebook or Twitter feeds for players who log into the game through the respective platforms. In addition, Fakey provides an anonymous mode for those who do not wish to login via Facebook or Twitter. Our goal was not to clone the Facebook or Twitter UIs (which are different from each other and change over time), but to present a generic conceptual replica of a social media news feed in which players are exposed to articles from diverse news sources and can take various actions related to these items.

Each round of play in Fakey presents players with a batch of ten news articles in the form of a news feed (see Section B in the Appendix). The game directions instruct players to `Share`, `Like`, `Fact Check`, or `Skip` the presented articles or use a `Hint` (see Section A in the Appendix). As shown in Figure 2, each item in the news feed consists of a corresponding photo, headline, description, social engagement metric, and action buttons.

Fakey helps players learn to pay attention to various signals, such as clickbait headlines, partisan or emotionally-charged language and images, and context. In addition, players can see the news source by using the `Hint` button. Players may choose to investigate the credibility of an article outside the game as well.

Fakey rewards players with points for each correct action and provides feedback for each article with which they interact (see Figure 3). The outcomes of these interactions are captured via two

Table 1. The points awarded for each action within Fakey, separated by mainstream and low-credibility articles.

	Mainstream		**Low-credibility**	
Action	Without hint	After hint	Without hint	After hint
`Share`	10	5	0	0
`Like`	8	4	2	1
`Fact Check`	4	2	10	5

scores — *Points* and *Skill* — displayed at the top of the news feed (see Figure 1). Points are awarded according to the scheme presented in Table 1. As Table 1 shows, players get the maximum number of points for sharing mainstream content or fact checking low-credibility articles. Using a hint prior to taking an action halves the number of points received for taking the same action without the use of the hint. Skill is the average of the points accumulated across game rounds. As such, the skill indicates whether a player is improving with each round. After finishing a round of ten articles, players can choose to continue to another round or check the leaderboard to compare their Points and Skill with those of top-performing players. Player actions within Fakey are recorded in an analytics database.

3.2 News Source Selection

Fakey uses the News[11] and Hoaxy[12] APIs to extract articles from mainstream and low-credibility sources, respectively. This approach follows the practice of analyzing content credibility at the domain (website) level rather than the article level. Such an approach circumvents the challenge of assessing the accuracy of individual news articles, which is infeasible at scale [8, 24, 38, 59, 65, 66]. Moreover, Shao et al. [65, 66] found that 82% of tweets that link to articles from low-credibility sources used by Fakey make claims that are classified as misinformation. The false positive rate (i.e., factually correct content) for these sources was found to be only 14%. Flagging or fact checking all articles from a source with a high rate of false claims is an appropriate action, even if fact checking confirms that a specific article from that source is accurate. Such an operation is analogous to how users may interact with low-credibility content on social media platforms.

For mainstream news, we manually selected 32 sources with a balance of moderate liberal, centrist, and moderate conservative views: *ABC News Australia, Al Jazeera English, Ars Technica, Associated Press, BBC News, Bloomberg, Business Insider, Buzzfeed, CNBC, CNN, Engadget, Financial Times, Fortune, Independent, Mashable, National Geographic, New Scientist, Newsweek, New York Magazine, Recode, Reuters, Techcrunch, The Economist, The Guardian, The New York Times, Next Web, Telegraph, Verge, The Wall Street Journal, The Washington Post, Time, and USA Today* [4]. We chose low-credibility sources based on flagging by various reputed fact-checking organizations, such as Snopes [65, 66]. We consider these sources as 'low credibility' because they tend to publish fake news, conspiracy theories, clickbait, rumors, junk science, and other questionable content.

Importantly, Fakey does not use a fixed set of labeled articles from mainstream and low-credibility sources. Rather, it shows *current* news articles selected from these sources (see Section B in the Appendix for an example showing articles presented within a round). Since articles within Fakey are retrieved in real time using public APIs, players always obtain fresh content, thus avoiding repeated exposure to an article. In fact, Fakey ensures that a player never encounters the same

[11] https://newsapi.org
[12] http://rapidapi.com/truthy/api/hoaxy

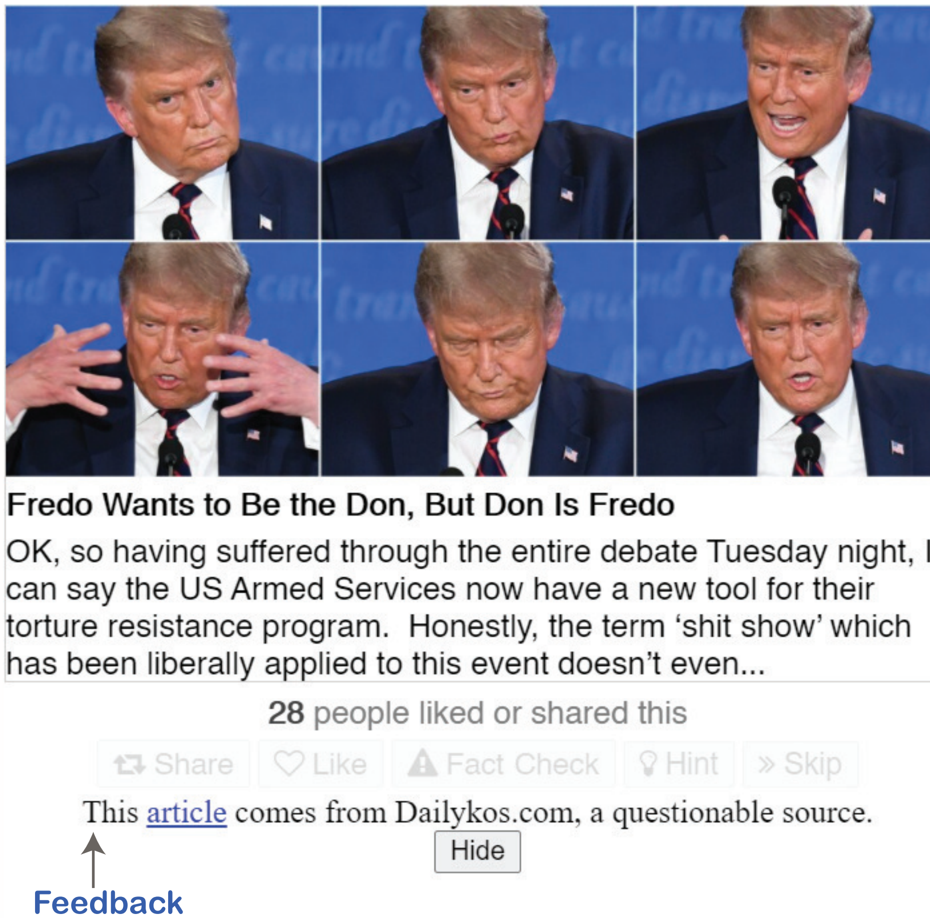


Fig. 3. Feedback provided by the game after the user performs an action on a news item.

article twice. Although it is conceivable that a player has previously come across an article outside of Fakey, the recency of articles presented within Fakey significantly decreases the likelihood of such prior knowledge. As a result, prior exposure to an article is likely to have minimal, if any, impact on player actions within Fakey that we analyzed in our study.

3.3 Social Engagement

For each article, Fakey displays a single metric for social engagement, i.e., the number of people who liked or shared the article. Having a single metric decreases the cognitive load for players and simplifies analysis compared to having distinct metrics for likes and shares. Engagement values η are randomly drawn from an approximately log-normal distribution with a maximum possible cutoff value of $\eta = 10^6$. The distribution is such that roughly 69% of the articles display engagement values $\eta > 10^2$, with roughly 3% showing values $\eta > 10^5$. Although the simulated engagement in the game is not drawn from empirical data, the metric numbers shown have a heavy tail similar to those typically observed on social media [79].

3.4 Game Mechanics

To select the ten articles shown to a player in each round, Fakey randomly picks five articles from mainstream and five from low-credibility sources, avoiding duplication within and across rounds and minimizing the selection of articles from the same source. For a given source, any article returned by the News or Hoaxy APIs is shown regardless of topic, without further filtering. Players interact with each article in the feed using one of the following action buttons:

3.4.1 Share: A social media user who shares an article or post participates in its spread. In the real world, users could share questionable content while stating that they do not endorse it. However, within the context of our research, when a player shares an article in the game, it means that the player endorses the article and wishes to share it with the world. Players are informed of this assumption in the instructions displayed before the start of the game (see Section A in the Appendix). Therefore, players who choose to share articles from low-credibility sources do not gain any points (see the `Share` row in Table 1). On the other hand, players receive the maximum number of points (i.e., 10 points) for sharing mainstream news.

3.4.2 Like: Liking an article on social media platforms is highly correlated with approving or endorsing its content [39]. Therefore, we assume that when a player likes an article in the game, it means that the player is endorsing the article. This assumption is mentioned in the initial game instructions (see Section A in the Appendix). Since liking is not as influential as sharing [55], players receive fewer points for liking a mainstream news article compared to the points received for sharing it (see Table 1). Further, players who like an article from a low-credibility source receive a minimal number of points (i.e., 2 points) compared to those awarded for liking a mainstream article (i.e., 8 points).

3.4.3 Fact Check: On many social media platforms (e.g., Facebook, Instagram, and Twitter), a user has the option to report or flag a potentially harmful article or post. The `Fact Check` button within Fakey mimics the mechanism of flagging suspicious content by seeking additional verification. As explained in the game instructions (see Section A in the Appendix), Fakey assumes that fact checking an article is a signal that the player does not trust the content. Therefore, fact checking a news article from a low-credibility source yields the maximum number of points (see the `Fact Check` row in Table 1). In contrast, fact checking a mainstream article is awarded relatively fewer points (i.e., 4 points).

3.4.4 Hint: When unsure, Fakey players can access a hint about the article. Clicking on the `Hint` button displays a link to the article's URL and mentions its source along with whether it is considered mainstream or low credibility. `Share`, `Like`, or `Fact Check` actions taken after accessing the hint are awarded half the number of points compared to the corresponding points that would have been received without consulting the hint (see Table 1).

3.4.5 Skip: As in any social media platform, the player can choose not to interact with an article, so Fakey provides a `Skip` button that does not impact the score in any way.

3.5 Deployment and Recruitment

Fakey was developed using MySQL, Python, and Django for the back end, and JavaScript, CSS, Vue, Quasar, and Apache Cordova for the front end. Fakey analytics capture player interaction with each article shown. Along with a timestamp, the recorded information includes whether the player shared, liked, or fact checked the article and whether a hint was used. In addition, the analytics record whether the article was retrieved from a mainstream or low-credibility source and the simulated social engagement metric displayed with the article.

Since the beginning of May 2018, Fakey has been deployed through a web interface and available within the iOS and Android app stores in the major English-speaking countries (Australia, Canada, United Kingdom, and United States). People from other countries can play the game only via the web interface. After release, Fakey was advertised through several channels. The researchers involved in creating Fakey announced the game on social media through institutional and personal accounts. In addition, Indiana University issued a press release through its media channels, and The Conversation research news portal featured an article about the game.[13] Further, Fakey was promoted at conferences, in keynote talks, and via word-of-mouth. It should be noted that the recruitment is not tied to a controlled experiment with systematic and pre-planned study enrollment efforts. Fakey remains accessible to anyone worldwide, and our research is based on observational analyses of naturalistic data. For the analyses in this paper, we saved a snapshot of the analytics database on November 26, 2019. The findings are therefore based on the analytics collected over 19 months between May 2018 and November 2019.[14]

3.6 Game Analytics

To respect player privacy and maximize participation, Fakey does not collect any user information apart from game analytics. Consequently, we did not have any means to reach out to players to compensate them. To provide some context regarding the findings presented in Section 4, we describe high-level usage data collected from Google Analytics. Most players played the game via their mobile devices (74%), which included tablets (4%). The rest played from desktop or laptop computers (26%). Fakey was accessed the most from Japan (~150k users). However, we excluded these game sessions because their average duration is only 14 seconds, suggesting that most of them did not involve any game-playing activity. Similarly, we excluded 533 sessions from South Korea with average duration of 4 seconds. Considering only players from countries with an average session duration of more than 40 seconds, most of the data reported in the paper comes from those in the United States (78%), Australia (8%), United Kingdom (4%), Canada (3.7%), Germany (3.6%), and Bulgaria (2.4%). These analytics reflect that the game was advertised heavily in the United States, through a US university, and in US-based conferences.

3.7 Semi-Structured Interviews

To verify and refine our understanding of how players interact with Fakey, we conducted a separate small-scale follow-up study in which participants played the game. We used a semi-structured protocol (see Section C in the Appendix) to interview the participants before and after they played the game. Six participants from six states in the United States (4 women, 2 men) with ages ranging from 19 to 46 (mean 28) were recruited via Reddit and Indiana University's classifieds board. The interview sessions lasted 38 minutes on average. We compensated each participant with $5 gift certificate to Amazon.com. We determined this incentive based on the minimum wage of US $7.25 per hour in our state.

4 FINDINGS

We aggregated the sessions of the same person based on login information (for players who logged in via Facebook or Twitter) or a cookie (for anonymous players). While this approach cannot mark all Fakey sessions of a single person with complete certainty, any disparity should be within the small margin of error typical of analytics-based user matching. With this caveat, the game session

[13] https://theconversation.com/misinformation-and-biases-infect-social-media-both-intentionally-and-accidentally-97148

[14] The data is available at: https://doi.org/10.7910/DVN/OPMIS4

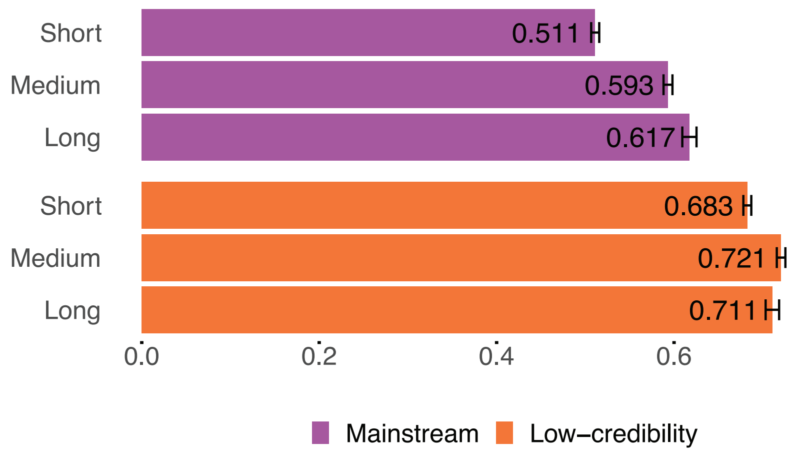


Fig. 4. Mean accuracy for the Short, Medium, and Long player groups, split by interactions with mainstream and low-credibility content.

data we analyzed involved 8,608 unique users who interacted with 119,166 news articles, distributed roughly equally between mainstream and low-credibility news sources.

For the present analysis, we consider clicking on the `Share` and `Like` buttons for articles from mainstream news sources along with clicking on the `Fact Check` button for articles from low-credibility sources as the most 'correct' actions. We first analyzed the `Share` and `Like` actions separately and found that these actions follow the same trends for all analyses. Therefore, we simplify the presentation of results by aggregating the `Share` and `Like` actions.

Kolmogorov-Smirnov tests showed that most of the data is not normally distributed ($p < 0.05$). Therefore, all statistical testing for differences was conducted using the non-parametric Kruskal-Wallis test, unless specified otherwise. For statistically significant differences, we further conducted post hoc pairwise Mann-Whitney tests with Bonferroni correction for multiple testing. We report effect sizes calculated by dividing the Z-statistic extracted from the pairwise comparisons by the square root of the sample size n, i.e., $Z/\sqrt{n}$ [19].

4.1 Effect of Intervention

To investigate the effect of session length, we grouped players into three bins: 'Short' for those who played a single round of the game (i.e., 1–10 articles); 'Medium' for those who played between 2 and 3 rounds (i.e., 11–30 articles); and 'Long' for those who played more than 3 rounds (i.e., more than 30 articles). For each player in these groups, we calculated the accuracy by adding the number of interactions in which they shared or liked mainstream articles with the number of interactions in which they fact-checked low-credibility articles and dividing the sum by the total number of articles. We did not find any statistical differences in the mean accuracy values across these player groups (Short: 0.636; Medium: 0.657; Long: 0.665; $p > 0.05$).

To understand the effectiveness of the game intervention, we investigated player interactions with mainstream and low-credibility content separately across the same player groups mentioned above (i.e., Short, Medium, Long). For each player we calculated: (1) the accuracy of recognizing mainstream content by dividing the number of `Share` and `Like` actions by the total number of interactions with mainstream content and (2) the accuracy of recognizing low-credibility content

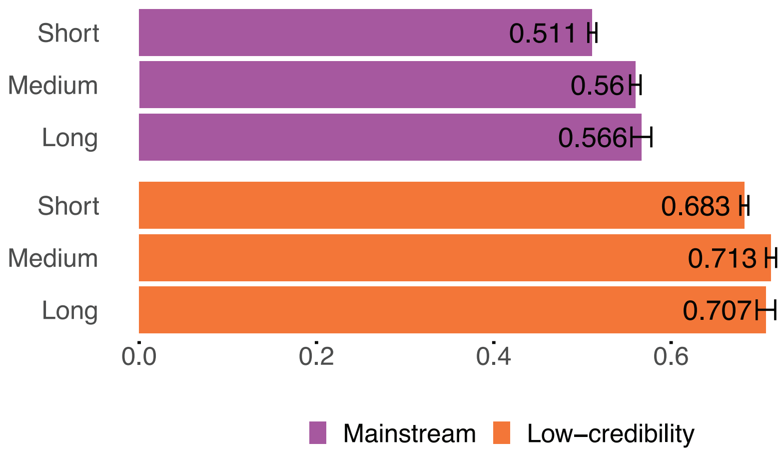


Fig. 5. Mean accuracy for the first round for the Short, Medium, and Long player groups, split by interactions with mainstream and low-credibility content.

by dividing the number of `Fact Check` actions by the number of interactions with low-credibility content (see Figure 4).

Kruskal-Wallis tests revealed statistically significant effects across player groups for mainstream articles ($\chi^2(2) = 79.47$, $p < 0.001$) as well as low-credibility articles ($\chi^2(2) = 24.91$, $p < 0.001$). We conducted post hoc Mann-Whitney tests to determine the specific comparisons that were statistically significant. We found statistically significant improvement when comparing the Short player group with Medium and Long player groups for mainstream content (Short-Medium: effect size = 0.081, $p < 0.001$; Short-Long: effect size = 0.075, $p < 0.001$) as well as low-credibility articles (Short-Medium: effect size = 0.037, $p < 0.01$; Short-Long: effect size = 0.053, $p < 0.001$). However, the effect size is higher for mainstream articles than that for low-credibility content. Comparing Medium and Long player groups revealed no statistically significant differences for either type of content. These results indicate that those who play more than one round are better at recognizing whether an article comes from a mainstream or low-credibility source.

The improvement in accuracy observed above could be attributed to self-selection (i.e., those who play multiple rounds having a predisposition to distinguish between mainstream and low-credibility content) or learning (i.e., those who play multiple rounds improving their news literacy because of the game). We conducted two further analyses aimed at disentangling the two factors. We first compared the accuracy of recognizing mainstream and low-credibility news for the first round (i.e., the first ten articles) for the three player groups defined above (see Figure 5). Kruskal-Wallis tests revealed statistically significant effects across player groups for mainstream articles ($\chi^2(2) = 23.00$, $p < 0.001$), but not for low-credibility ones ($\chi^2(2) = 2.58$, $p > 0.05$). We conducted post hoc Mann-Whitney tests to determine the specific comparisons that were statistically significant. We found statistically significant improvement when comparing the Short player group with Medium and Long player groups (Short-Medium: effect size = 0.046, $p < 0.001$; Short-Long: effect size = 0.038, $p < 0.001$). These results show that those who played for only a single round were worse at recognizing mainstream content during that round compared to the first-round performance of the others who continued on to play further rounds of the game. The lower performance could potentially have led these players (i.e., those in the Short group) to quit playing, thus indicating that self-selection might be a factor influencing the observed results.

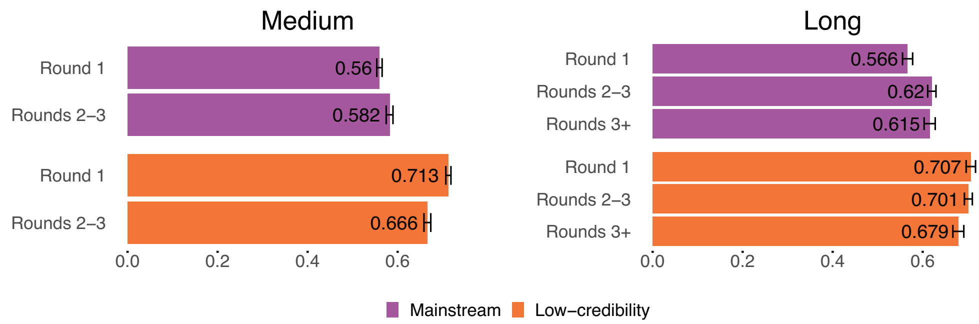


Fig. 6. Mean accuracy for the Medium and Long player groups binned by game rounds: Round 1, Rounds 2–3, and Rounds 3+.

To understand whether those who proceeded to play more than one round experienced a learning effect, we analyzed the Medium and Long player groups in detail. Within both of these player groups, we binned interactions into those representing the first round, those included in rounds 2 and 3, and those for rounds 4 and higher. Since those in the Medium player group played either 2 or 3 rounds, that group contained two bins, while the Long player group contained three bins because players in this group played more than 3 rounds. We then calculated the accuracy of recognizing mainstream and low-credibility articles for each bin within the Medium and Long player groups (see Figure 6).

For mainstream content, we found statistically significant *improvements* for Medium (Round 1 bin vs. Rounds 2–3 bin: effect size = 0.095, $p < 0.001$) as well as Long player groups (Round 1 bin vs. Rounds 2–3 bin: effect size = 0.195, $p < 0.001$; Round 1 bin vs. Rounds 3+ bin: effect size = 0.158, $p < 0.001$). On the other hand, for low-credibility content, we found a statistically significant *decrease* in accuracy for the Medium player group (Round 1 bin vs. Rounds 2–3 bin: effect size = 0.061, $p < 0.001$). Those who play up to three rounds of Fakey `Share` or `Like` more articles from mainstream as well as low-credibility sources. We did not find a corresponding statistically significant decrease in accuracy for players in the Long group, i.e., those who play more than three rounds learn to `Share` or `Like` more mainstream content, but avoid a similar increase in sharing or liking low-credibility articles.

The above results address our research question regarding the learning effect of Fakey: playing the game for more than three rounds improves the recognition of mainstream articles, without affecting the detection of low-credibility content. Yet, the fact that players do not get better at recognizing questionable articles underscores the challenge of quickly dismissing such content as fake; recall that the content from low-credibility sources may not always be misinformation. Further game mechanisms should be explored to disincentivize sharing or liking dubious content within the game.

4.2 Role of Hints

Across the dataset, players used the `Hint` button 6,676 times, i.e., in 5.6% of all recorded interactions. Players used the hint slightly more when dealing with mainstream articles (3,622 times, 54%) than when deciding on low-credibility content (3,054 times, 46%). We found that the hint feature was used similarly by all three player groups (Short: 2,212, 33%; Medium: 2,294, 34%; Long: 2,170, 33%). As observed in Figure 7, in more than 70% of the cases, the hint helped players subsequently take the correct action (i.e., sharing or liking mainstream content or fact checking low-credibility content).

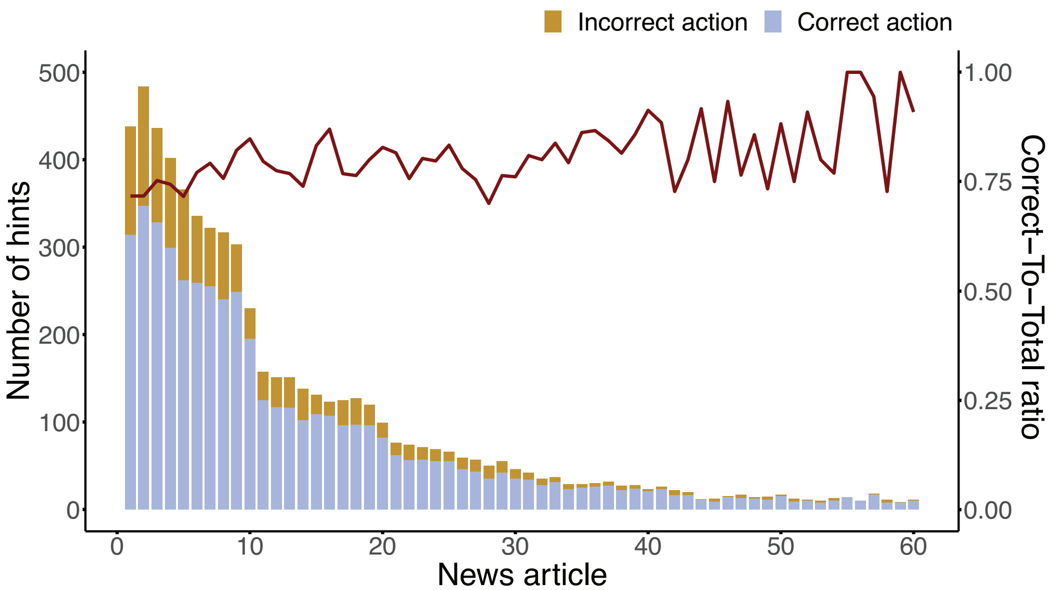


Fig. 7. Hints used during each article interaction within Fakey. The bars indicate the number of hints sought during the corresponding article interaction, with the light blue portion signifying the number of correct post-hint actions and the orange portion depicting the number of incorrect post-hint actions. The maroon horizontal line plots the ratio of correct post-hint actions to the number of hints for each article interaction.

Figure 7 further shows that the ratio of correct post-hint actions to the number of hints increases as players interact more with the game. We compared the respective ratios of correct post-hint actions to the number of hints for Round 1, Rounds 2–3, and Rounds 3+. Since the data is normally distributed ($p > 0.05$ using a Kolmogorov-Smirnov test), we conducted a one-way ANOVA that showed statistically significant differences across the three groups ($F(2, 57) = 9.24$, $p < 0.001$). Post hoc pairwise comparisons via the Tukey test found that the ratios of correct post-hint actions for Rounds 3+ are statistically significantly higher than those for Round 1 and Rounds 2–3 ($p < 0.01$). These differences suggest that the utility of the hint increases as players play longer.

Although the hint feature was used sporadically, it was highly effective when invoked and its utility increased with longer game play.

4.3 User Experience

Semi-structured interviews confirmed that players understand the purpose of Fakey and find the game useful. Interviewees described the game as simple, fun, and educational. Importantly, we confirmed that players understand the game UI and the action buttons as intended, especially after playing an initial round. In particular, interviewees equated the `Share` and `Like` buttons to the corresponding social media features and understood the combined number of shares and likes as indicating a metric related to the engagement of other people with the article.

> *"I think it's similar to social media. The more people that share or like something, the more likely it is to go viral or that it's probably more credible if it's shared that much."* – (P6, 20, Female)

Interviewees reported using all or most of the pieces of information within the game (i.e., photo, headline, social engagement metric, and, sometimes, the hint) to discern whether the presented article was high- or low-credibility.

> *"The photos helped a lot [in judging credibility]. But I also looked at other things, like the headline, like how it was written... if it sounds like something that could be true."* – (P4, 36, Male)

Further, as intended, interviewees reported using the `Fact Check` button when they were unsure about an article's credibility and desired additional fact-checking information. As P3 (19, Male) said, the `Fact Check` button was expected to help in *"confirming whether the claim is true or false."*

The interviews confirmed that Fakey seems to meet its goals of increasing awareness about news content consumed on social media and training players to be vigilant regarding misinformation and questionable content.

> *"It [Fakey] gave me an idea that I should be careful. When I'm reading a headline or an article, I should double check and make sure that what I'm reading is coming from a trusted source. The game definitely gave me a perception of how headlines and pictures can be very deceiving and very misleading when it comes to reading information online."* – (P5, 24, Female)

Interviewees specifically mentioned learning effects even when they felt that they did not perform especially well.

> *"I want to play it again to get better at those skills. I think I like the fact that it's testing your ability to know what's real and what's not. If that's something people have an interest in, then they should play this game."* – (P6, 20, Female)

5 DISCUSSION AND IMPLICATIONS

We discuss the implications of the main results related to our research question, followed by the limitations of our approach and data. We then offer promising directions for future research.

5.1 Priming Effect

Our results show that priming people to fact check makes them more suspicious of mainstream and low-credibility content alike. The priming leads people to dispute half of the mainstream and over 70% of the low-credibility content that they encounter in the game. These rates highlight that more than a quarter of questionable content still sneaks through without verification, even when people are explicitly primed to look for it. At the same time, people become overly suspicious of legitimate content. Our findings support the need for news consumers to pause and consider the credibility of headlines [6, 14, 58], engage in civic online reasoning [44], and develop critical thinking [42, 45]. However, it is important that such interventions not create excessive suspicion of legitimate content. Finding the right balance between excessive suspicion and insufficient scrutiny is an important challenge for future research.

One of the reasons for the inability of people to spot questionable content is the UI used to display news on social media. Plenty of research has shown evidence of the effectiveness of credibility indicators (e.g., [57, 58, 84]). Further empirical work is needed to explore how such indicators could be presented within the UI and the information they should contain.

5.2 Learning

We found that repeated play improves the ability to recognize mainstream content without affecting the ability to recognize questionable content. Further research is needed to determine the amount of game play required to achieve sufficiently high proficiency in separating legitimate and questionable content and the length of time the acquired skill is retained.

Games like Fakey could be useful if administered as an intervention to social media users. For instance, social media platforms could conduct regular drills (akin to 'phishing drills' used in organizations for employee security training) wherein users are trained using examples of debunked articles. Such games could even be integrated into media literacy curricula in schools.

5.3 Player Retention

When considering only the results of the first round played, we found that players who stop playing the game within a single round perform worse than those who play more than one round. Semi-structured interviews indicated that some players may find the point scheme to be less rewarding than expected.

Since our results show that players who continue playing improve their performance, the interview findings imply that further research is needed to find techniques that can retain players who do not perform well in initial rounds. Player retention could be addressed by employing techniques from research on persuasion [16, 47]. For instance, players could be encouraged to use the `Hint` button right after performing the first incorrect action or could be shown encouraging messages when they make a mistake [17, 48]. If these techniques are evaluated to be successful in retaining players who do not perform well in the initial rounds, then the game has the potential of becoming more effective in its goal of training users to become more vigilant when consuming online content.

5.4 Limitations and Future Work

The research reported in this paper focuses on user interaction elements within Fakey. Of course, various other factors influence the spread of online misinformation. For instance, knowing about education, demographics, and/or political affiliation of players might provide further insight. However, to respect privacy, Fakey does not collect any player information except game analytics. For instance, we cannot associate a country of origin with a specific player session because Fakey analytics are anonymous and country information is obtained separately from Google analytics. Similarly, we have no data to comment on representativeness of players from each country.

The types of content that are more likely to be influenced by engagement metrics would further improve our understanding of misinformation spread. However, to limit player cognitive burden and simplify analysis, Fakey currently displays a single social engagement metric that combines the number of shares and likes. During the semi-structured interviews, none of the interviewees reported that this design choice impacted their understanding of the social engagement metric. However, we acknowledge that social media platforms do not show social engagement in the form of such a combined metric and might show additional elements such as comments.

In each round, Fakey shows five news articles each from mainstream and low-credibility sources. This is likely a higher proportion of misinformation compared to that experienced by most social media users in the wild. As a result, the game setting might make players more suspicious within the game compared to the real world, increasing fact-checking rates. However, this effect applies to both types of sources, so comparative results are unaffected by the priming effect.

News sources are typically not displayed prominently on social media news feeds, such as those of Facebook and Twitter. As a result, social media users are vulnerable to content from questionable sources to a greater extent. A goal of Fakey is to help social media users become aware of this vulnerability. Therefore, Fakey interface mimics the lack of emphasis on sources in real-world social-media feeds. However, when a player takes an incorrect action or asks for a hint, Fakey provides feedback that includes the article source. In this way, players have a chance to realize the importance of the source and transfer that knowledge to the real world. Our analysis indeed suggests that hints about article sources help improve news literacy. Future work could evaluate variations in the presentation of source information to achieve optimal learning.

By emulating the relevant interface elements of popular social media platforms, such as Facebook and Twitter, Fakey avoids the ethical concerns of manipulating content in real-world news feeds [35]. However, a simulated social media environment has clear limitations as it differs from the actual platform. For instance, Fakey minimizes the cognitive load of players by capturing only the `Like` and `Share` actions, which were the earliest ones on social media platforms and, as such, are the most common across platforms and most familiar to users.

In addition to an inability to decide whether to share or like an article, skipping could be attributed to other reasons, such as a lack of interest or a lack of attention. Unfortunately, the game platform does not collect data about the `Skip` action, so this information is not available for analysis. Since Fakey is designed to simulate a social media news feed in which users scroll through items without necessarily interacting with each one, we did not feel the need to capture explicit `Skip` actions. Due to this design choice, the data that we are able to analyze is limited to the cases when a player performed an action other than `Skip`. Investigating when, how, and why people decide to skip articles is an interesting avenue for future research.

6 CONCLUSION

Game-based approaches can be useful to help people improve news literacy and avoid falling victim to misinformation. Given that social media is a primary vehicle for misinformation spread, such approaches must be effective in countering exposure to questionable content in a social media context. To that end, we designed a game that emulates the experience of a real-world social media news feed to train people to verify questionable content and avoid endorsing and spreading it to their friends. We confirmed that the game can enhance information evaluation skills and increase digital literacy. Our research underscores that countering online misinformation is a multi-faceted challenge that requires paying attention not just to the content and the actors but also to the UX and UIs of the platforms on which the content is disseminated.

ACKNOWLEDGMENTS

We are grateful to Alessandro Flammini for helpful discussion and Chengcheng Shao and Clayton A. Davis for technical support during platform development. We thank the anonymous players of Fakey and the participants of the interview study. This research is supported in part by the Democracy Fund, Craig Newmark Philanthropies, and the Knight Foundation. The contents of the paper are the work of the authors and do not reflect the views of the sponsors.

REFERENCES

[1] Michelle A. Amazeen, Emily Thorson, Ashley Muddiman, and Lucas Graves. 2018. Correcting political and consumer Misperceptions: The effectiveness and effects of rating scale versus contextual correction formats. *Journalism & Mass Communication Quarterly* 95, 1 (2018), 28–48. https://doi.org/10.1177/1077699016678186

[2] Ahmer Arif, Leo Graiden Stewart, and Kate Starbird. 2018. Acting the part: Examining information operations within #BlackLivesMatter discourse. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 20 (Nov. 2018), 27 pages. https://doi.org/10.1145/3274289

[3] Patricia Aufderheide. 1993. Media literacy. A report of the National Leadership Conference on Media Literacy. *Education Resources Information Center* ED365294 (1993), 44 pages. https://files.eric.ed.gov/fulltext/ED365294.pdf

[4] Mihai Avram, Nicholas Micallef, Sameer Patil, and Filippo Menczer. 2020. Exposure to social engagement metrics increases vulnerability to misinformation. *The Harvard Kennedy School Misinformation Review* 1, 5 (2020), 1–11. https://doi.org/10.37016/mr-2020-033

[5] Melisa Basol, Jon Roozenbeek, and Sander van der Linden. 2020. Good news about bad news: Gamified inoculation boosts confidence and cognitive immunity against fake news. *Journal of cognition* 3, 1 (2020), 1–9. https://doi.org/10.5334/joc.91

[6] Md. Momen Bhuiyan, Kexin Zhang, Kelsey Vick, Michael A. Horning, and Tanushree Mitra. 2018. FeedReflect: A tool for nudging users to assess news credibility on Twitter. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Jersey City, NJ, USA) *(CSCW '18)*. Association for Computing Machinery, New York, NY, USA, 205–208. https://doi.org/10.1145/3272973.3274056

[7] Leticia Bode and Emily K. Vraga. 2015. In related news, that was wrong: The correction of misinformation through related stories functionality in social media. *Journal of Communication* 65, 4 (Jun 2015), 619–638. https://doi.org/10.1111/jcom.12166

[8] Alexandre Bovet and Hernán A. Makse. 2019. Influence of fake news in Twitter during the 2016 US presidential election. *Nature Communications* 10, 1 (2019), 1–14. https://doi.org/10.1038/s41467-018-07761-2

[9] Ceren Budak, Divyakant Agrawal, and Amr El Abbadi. 2011. Limiting the spread of misinformation in social networks. In *Proceedings of the 20th International Conference on World Wide Web* (Hyderabad, India) *(WWW '11)*. Association for Computing Machinery, New York, NY, USA, 665–674. https://doi.org/10.1145/1963405.1963499

[10] Robert Chesney and Danielle Citron. 2019. Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. *Foreign Affairs* 98, 1 (2019).

[11] Katherine Clayton, Spencer Blair, Jonathan A. Busam, Samuel Forstner, John Glance, Guy Green, Anna Kawata, Akhila Kovvuri, Jonathan Martin, Evan Morgan, Morgan Sandhu, Rachel Sang, Rachel Scholz-Bright, Austin T. Welch, Andrew G. Wolff, Amanda Zhou, and Brendan Nyhan. 2020. Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. *Political Behavior* 42, 4 (2020), 1073–1095. https://doi.org/10.1007/s11109-019-09533-0

[12] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. 2016. The spreading of misinformation online. *Proceedings of the National Academy of Sciences* 113, 3 (2016), 554–559. https://doi.org/10.1073/pnas.1517441113

[13] Nicholas Dias, Gordon Pennycook, and David G. Rand. 2020. Emphasizing publishers does not effectively reduce susceptibility to misinformation on social media. *The Harvard Kennedy School Misinformation Review* 1, 1 (2020), 1–12. https://doi.org/10.37016/mr-2020-001

[14] Lisa Fazio. 2020. Pausing to consider why a headline is true or false can help reduce the sharing of false news. *The Harvard Kennedy School Misinformation Review* 1, 2 (2020), 1–8. https://doi.org/10.37016/mr-2020-009

[15] Seth Flaxman, Sharad Goel, and Justin M. Rao. 2016. Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly* 80, S1 (Mar 2016), 298–320. https://doi.org/10.1093/poq/nfw006

[16] Brian Jeffrey Fogg. 2002. *Persuasive technology: Using computers to change what we think and do* (1 ed.). Number December. Morgan Kaufmann, New York, NY, USA. https://doi.org/10.1145/764008.763957

[17] Alain Forget, Sonia Chiasson, and Robert Biddle. 2009. Lessons from brain age on persuasion for computer security. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI EA '09)*. Association for Computing Machinery, New York, NY, USA, 4435–4440. https://doi.org/10.1145/1520340.1520679

[18] Adrien Friggeri, Lada Adamic, Dean Eckles, and Justin Cheng. 2014. Rumor cascades. *Proceedings of the International AAAI Conference on Web and Social Media* 8, 1 (2014). https://ojs.aaai.org/index.php/ICWSM/article/view/14559

[19] Catherine O. Fritz, Peter E. Morris, and Jennifer J. Richler. 2012. Effect size estimates: Current use, calculations, and interpretation. *Journal of experimental psychology: General* 141, 1 (2012), 2–18. https://doi.org/10.1037/a0024338

[20] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. 2017. Reducing controversy by connecting opposing views. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining* (Cambridge, United Kingdom) *(WSDM '17)*. Association for Computing Machinery, New York, NY, USA, 81–90. https://doi.org/10.1145/3018661.3018703

[21] R. Kelly Garrett and Brian E. Weeks. 2013. *The promise and peril of real-time corrections to political misperceptions.* Association for Computing Machinery, New York, NY, USA, 1047–1058. https://doi.org/10.1145/2441776.2441895

[22] Maria Glenski, Corey Pennycuff, and Tim Weninger. 2017. Consumers and curators: Browsing and voting patterns on Reddit. *IEEE Transactions on Computational Social Systems* 4, 4 (2017), 196–206. https://doi.org/10.1109/TCSS.2017.2742242

[23] Samuel Greengard. 2020. Will deepfakes do deep damage? *Commun. ACM* 63, 1 (2020), 17–19.

[24] Nir Grinberg, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. 2019. Fake news on Twitter during the 2016 US presidential election. *Science* 363, 6425 (2019), 374–378. https://doi.org/10.1126/science.aau2706

[25] Andrew Guess, Jonathan Nagler, and Joshua Tucker. 2019. Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances* 5, 1 (2019), 1–8. https://doi.org/10.1126/sciadv.aau4586

[26] Andrew Guess, Brendan Nyhan, and Jason Reifler. 2018. Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. *European Research Council* 9, 3 (2018), 1–14.

[27] Jimmeka J. Guillory and Lisa Geraci. 2013. Correcting erroneous inferences in memory: The role of source credibility. *Journal of Applied Research in Memory and Cognition* 2, 4 (2013), 201–209. https://doi.org/10.1016/j.jarmac.2013.10.001

[28] Aditi Gupta, Hemank Lamba, Ponnurangam Kumaraguru, and Anupam Joshi. 2013. Faking Sandy: Characterizing and identifying fake images on Twitter during Hurricane Sandy. In *Proceedings of the 22nd International Conference on World Wide Web* (Rio de Janeiro, Brazil) *(WWW '13 Companion)*. Association for Computing Machinery, New York, NY, USA, 729–736. https://doi.org/10.1145/2487788.2488033

[29] Michael Hameleers, Thomas E. Powell, Toni G. L. A. Van Der Meer, and Lieke Bos. 2020. A picture paints a thousand lies? The effects and mechanisms of multimodal disinformation and rebuttals disseminated via social media. *Political Communication* 37, 2 (2020), 281–301. https://doi.org/10.1080/10584609.2019.1674979

[30] Michael Hameleers and Toni G. L. A. van der Meer. 2020. Misinformation and polarization in a high-Choice media environment: How effective are political fact-Checkers? *Communication Research* 47, 2 (2020), 227–250. https://doi.org/10.1177/0093650218819671

[31] Jane Im, Sonali Tandon, Eshwar Chandrasekharan, Taylor Denby, and Eric Gilbert. 2020. Synthesized social signals: Computationally-derived social signals from account histories. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376383

[32] Dean Jackson. 2017. Issue brief: Distinguishing disinformation from propaganda, misinformation, and "fake news". *National Endowment for Democracy* (2017). https://www.ned.org/issue-brief-distinguishing-disinformation-from-propaganda-misinformation-and-fake-news

[33] Maurice Jakesch, Moran Koren, Anna Evtushenko, and Mor Naaman. 2019. The role of source and expressive responding in political news evaluation. Presented at Computation and Journalism Symposium. https://people.jacobs.cornell.edu/mor/publications/thegoods/jakesch_cplusj19__trustnews.pdf

[34] Antonis Kalogeropoulos, Richard Fletcher, and Rasmus Kleis Nielsen. 2019. News brand attribution in distributed environments: Do people know where they get their news? *New Media & Society* 21, 3 (2019), 583–601. https://doi.org/10.1177/1461444818801313

[35] Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 111, 24 (2014), 8788–8790. https://doi.org/10.1073/pnas.1320040111

[36] Srijan Kumar and Neil Shah. 2018. False information on Web and social media: A survey. In *Social Media Analytics: Advances and Applications*. CRC Press.

[37] Srijan Kumar, Robert West, and Jure Leskovec. 2016. Disinformation on the web: Impact, characteristics, and detection of Wikipedia hoaxes. In *Proceedings of the 25th International Conference on World Wide Web* (Montréal, Québec, Canada) *(WWW '16)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 591–602. https://doi.org/10.1145/2872427.2883085

[38] David M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, and Jonathan L. Zittrain. 2018. The science of fake news. *Science* 359, 6380 (2018), 1094–1096. https://doi.org/10.1126/science.aao2998

[39] Shu-Yueh Lee, Sara Steffes Hansen, and Jin Kyun Lee. 2016. What makes us click "like" on Facebook? Examining psychological, technological, and motivational factors on virtual endorsement. *Computer Communications* 73 (2016), 332–341. https://doi.org/10.1016/j.comcom.2015.08.002 Online Social Networks.

[40] Stephan Lewandowsky, Ullrich K. H. Ecker, and John Cook. 2017. Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of Applied Research in Memory and Cognition* 6, 4 (2017), 353–369. https://doi.org/10.1016/j.jarmac.2017.07.008

[41] Sonia Livingstone. 2004. Media literacy and the challenge of new information and communication technologies. *The Communication Review* 7, 1 (2004), 3–14. https://doi.org/10.1080/10714420490280152

[42] Lauren Lutzke, Caitlin Drummond, Paul Slovic, and Joseph Árvai. 2019. Priming critical thinking: Simple interventions limit the influence of fake news about climate change on Facebook. *Global Environmental Change* 58, 101964 (2019), 1–8. https://doi.org/10.1016/j.gloenvcha.2019.101964

[43] Drew B. Margolin, Aniko Hannak, and Ingmar Weber. 2018. Political fact-Checking on Twitter: When do corrections have an effect? *Political Communication* 35, 2 (2018), 196–219. https://doi.org/10.1080/10584609.2017.1334018

[44] Sarah McGrew. 2020. Learning to evaluate: An intervention in civic online reasoning. *Computers & Education* 145, 103711 (2020), 1–13. https://doi.org/10.1016/j.compedu.2019.103711

[45] Sarah McGrew, Joel Breakstone, Teresa Ortega, Mark Smith, and Sam Wineburg. 2018. Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & Research in Social Education* 46, 2 (2018), 165–193. https://doi.org/10.1080/00933104.2017.1416320

[46] Solomon Messing and Sean J. Westwood. 2014. Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communication Research* 41, 8 (2014), 1042–1063. https://doi.org/10.1177/0093650212466406

[47] Nicholas Micallef and Nalin Asanka Gamagedara Arachchilage. 2017. A gamified approach to improve users' memorability of fall-back authentication. In *Thirteenth Symposium on Usable Privacy and Security* (Santa Clara, CA) *(SOUPS 2017)*. USENIX Association. https://www.usenix.org/conference/soups2017/workshop-program/way2017/micallef

[48] Nicholas Micallef and Nalin Asanka Gamagedara Arachchilage. 2018. Security questions education: Exploring gamified features and functionalities. *Information & Computer Security* 26, 3 (2018), 365–378. https://doi.org/10.1108/ICS-03-2018-0033

[49] Nicholas Micallef, Bing He, Srijan Kumar, Mustaque Ahamad, and Nasir Memon. 2020. The role of the crowd in countering misinformation: A case study of the COVID-19 infodemic. In *2020 IEEE International Conference on Big Data (BigData 2020)*. IEEE. https://arxiv.org/abs/2011.05773

[50] Vidya Narayanan, Vlad Barash, John Kelly, Bence Kollanyi, Lisa-Maria Neudert, and Philip N. Howard. 2018. Polarization, partisanship and junk news consumption over social media in the US. arXiv:1803.01845 [cs.SI]

[51] Raymond S. Nickerson. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* 2, 2 (1998), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

[52] Thomas Nygren, Fredrik Brounéus, and Göran Svensson. 2019. Diversity and credibility in young people's news feeds: A foundation for teaching and learning citizenship in a digital era. *Journal of Social Science Education* 18, 2 Summer 2019 (2019), 87–109. https://doi.org/10.4119/jsse-917

[53] Brendan Nyhan, Ethan Porter, Jason Reifler, and Thomas J. Wood. 2020. Taking fact-Checks literally but not seriously? The effects of journalistic fact-checking on factual beliefs and candidate favorability. *Political Behavior* 42, 3 (2020), 939–960. https://doi.org/10.1007/s11109-019-09528-x

[54] Brendan Nyhan and Jason Reifler. 2010. When corrections fail: The persistence of political misperceptions. *Political Behavior* 32, 2 (2010), 303–330. https://doi.org/10.1007/s11109-010-9112-2

[55] Anne Oeldorf-Hirsch and S. Shyam Sundar. 2015. Posting, commenting, and tagging: Effects of sharing news stories on Facebook. *Computers in Human Behavior* 44 (2015), 240–249. https://doi.org/10.1016/j.chb.2014.11.024

[56] Pinar Ozturk, Huaye Li, and Yasuaki Sakamoto. 2015. Combating rumor spread on social media: The effectiveness of refutation and warning. In *2015 48th Hawaii International Conference on System Sciences*. 2406–2414. https://doi.org/10.1109/HICSS.2015.288

[57] Gordon Pennycook, Adam Bear, Evan T. Collins, and David G. Rand. 2020. The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management Science* 66, 11 (2020), 4944–4957. https://doi.org/10.1287/mnsc.2019.3478

[58] Gordon Pennycook, Jonathon McPhetres, Yunhao Zhang, Jackson G. Lu, and David G. Rand. 2020. Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science* 31, 7 (2020), 770–780. https://doi.org/10.1177/0956797620939054

[59] Gordon Pennycook and David G. Rand. 2019. Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences* 116, 7 (2019), 2521–2526. https://doi.org/10.1073/pnas.1806781116

[60] Gordon Pennycook and David G. Rand. 2020. Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality* 88, 2 (2020), 185–200. https://doi.org/10.1111/jopy.12476

[61] Verónica Pérez-Rosas, Bennett Kleinberg, Alexandra Lefevre, and Rada Mihalcea. 2018. Automatic detection of fake news. In *Proceedings of the 27th International Conference on Computational Linguistics*. Association for Computational Linguistics, Santa Fe, New Mexico, USA, 3391–3401. https://www.aclweb.org/anthology/C18-1287

[62] Walter Quattrociocchi, Antonio Scala, and Cass R. Sunstein. 2016. Echo chambers on Facebook. *Social Science Research Network* (2016), 1–15. https://doi.org/10.2139/ssrn.2795110

[63] Jon Roozenbeek and Sander van der Linden. 2019. The fake news game: Actively inoculating against the risk of misinformation. *Journal of Risk Research* 22, 5 (2019), 570–580. https://doi.org/10.1080/13669877.2018.1443491

[64] Jon Roozenbeek and Sander van der Linden. 2019. Fake news game confers psychological resistance against online misinformation. *Palgrave Communications* 5, 1, Article 65 (2019), 10 pages. https://doi.org/10.1057/s41599-019-0279-9

[65] Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini, and Filippo Menczer. 2018. The spread of low-credibility content by social bots. *Nature Communications* 9, 4787 (2018), 1–9. https://doi.org/10.1038/s41467-018-06930-7

[66] Chengcheng Shao, Pik-Mai Hui, Lei Wang, Xinwen Jiang, Alessandro Flammini, Filippo Menczer, and Giovanni Luca Ciampaglia. 2018. Anatomy of an online misinformation network. *PLOS ONE* 13, 4 (Apr 2018), 1–23. https://doi.org/10.1371/journal.pone.0196087

[67] Cuihua Shen, Mona Kasra, Wenjing Pan, Grace A. Bassett, Yining Malloch, and James F. O'Brien. 2019. Fake images: The effects of source, intermediary, and digital media literacy on contextual assessment of image credibility online. *New Media & Society* 21, 2 (2019), 438–463. https://doi.org/10.1177/1461444818799526

[68] Jieun Shin and Kjerstin Thorson. 2017. Partisan selective sharing: The biased diffusion of fact-checking messages on social media. *Journal of Communication* 67, 2 (Feb 2017), 233–255. https://doi.org/10.1111/jcom.12284

[69] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake news detection on social media: A data mining perspective. *SIGKDD Explor. Newsl.* 19, 1 (Sep 2017), 22–36. https://doi.org/10.1145/3137597.3137600

[70] Craig Silverman. 2015. Lies, damn lies, and viral content. *Tow Center for Digital Journalism, Columbia Journalism School* A Tow/Knight Report (2015). https://doi.org/10.7916/D8Q81RHH

[71] Craig Silverman. 2016. This analysis shows how viral fake election news stories outperformed real news on Facebook. *BuzzFeed News* (2016). https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook

[72] Kate Starbird. 2019. Disinformation's spread: Bots, trolls and all of us. *Nature* 571, 7766 (2019).

[73] Kate Starbird, Jim Maddock, Mania Orand, Peg Achterman, and Robert M. Mason. 2014. Rumors, false flags, and digital vigilantes: Misinformation on Twitter after the 2013 Boston marathon bombing. In *Proceedings of iConference 2014*. iSchools. https://doi.org/10.9776/14308

[74] David Sterrett, Dan Malato, Jennifer Benz, Liz Kantor, Trevor Tompson, Tom Rosenstiel, Jeff Sonderman, and Kevin Loker. 2019. Who shared it?: Deciding what news to trust on social media. *Digital Journalism* 7, 6 (2019), 783–801. https://doi.org/10.1080/21670811.2019.1623702

[75] Leo G. Stewart, Ahmer Arif, and Kate Starbird. 2018. Examining trolls and polarization with a retweet network. In *Proceedings of MIS2: Misinformation and Misbehavior Mining on the Web (Workshop held in conjunction with WSDM 2018)*. http://snap.stanford.edu/mis2/files/MIS2_paper_21.pdf

[76] Charles S. Taber and Milton Lodge. 2006. Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science* 50, 3 (2006), 755–769. https://doi.org/10.1111/j.1540-5907.2006.00214.x

[77] Emily Thorson. 2016. Belief echoes: The persistent effects of corrected misinformation. *Political Communication* 33, 3 (2016), 460–480. https://doi.org/10.1080/10584609.2015.1102187

[78] Sander van der Linden. 2017. Beating the hell out of fake news. *Ethical Record: Proceedings of the Conway Hall Ethical Society* 122, 6 (2017), 4–7. https://ssrn.com/abstract=3089590

[79] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science* 359, 6380 (2018), 1146–1151. https://doi.org/10.1126/science.aap9559

[80] Emily K. Vraga and Leticia Bode. 2017. Using expert sources to correct health misinformation in social media. *Science Communication* 39, 5 (2017), 621–645. https://doi.org/10.1177/1075547017731776

[81] Emily K. Vraga and Leticia Bode. 2018. I do not believe you: How providing a source corrects health misperceptions across social media platforms. *Information, Communication & Society* 21, 10 (2018), 1337–1353. https://doi.org/10.1080/1369118X.2017.1313883

[82] Emily K. Vraga, Leticia Bode, and Melissa Tully. 2020. Creating news literacy messages to enhance expert corrections of misinformation on Twitter. *Communication Research* (Jan 2020), 1–23. https://doi.org/10.1177/0093650219898094

[83] Thomas Wood and Ethan Porter. 2019. The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior* 41, 1 (2019), 135–163. https://doi.org/10.1007/s11109-018-9443-y

[84] Waheeb Yaqub, Otari Kakhidze, Morgan L. Brockman, Nasir Memon, and Sameer Patil. 2020. Effects of credibility indicators on social media news sharing intent. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376213

[85] Dannagal G. Young, Kathleen Hall Jamieson, Shannon Poulsen, and Abigail Goldring. 2018. Fact-checking effectiveness as a function of format and tone: Evaluating Factcheck.org and Flackcheck.org. *Journalism & Mass Communication Quarterly* 95, 1 (2018), 49–75. https://doi.org/10.1177/1077699017710453

A GAME INSTRUCTIONS

A.1 Introductory page

How to Play:

This game aims to teach media literacy and study how people interact with misinformation.

You will see a simulated news feed with articles like the one below: some coming from legitimate news sources and some from sites that typically publish false or misleading reports, clickbait headlines, conspiracy theories, junk science, and other types of misinformation.

Your goal is to support a healthy social media experience by promoting information from reliable sources and not from low-credibility sources.

A.2 Page 2

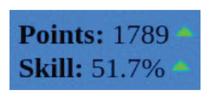
Inspect each article in the feed just as you would in your favorite social media service. Look at the image, headline, and description to decide whether it's credible, in which case, you may share or like it. Select `Fact Check` to indicate that you don't trust the article.

These actions are just simulations and will not affect your social media profile. Select `Hint` or `Skip` if you are not sure.

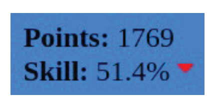
A.3 Page 3

The game will score each action you take. Your points describe how much experience you have. Your skill measures how good you are at promoting information from trustworthy sources and spotting articles from low-credibility sources.

You will get top points for sharing articles from credible sources and fact checking articles from suspicious sources.



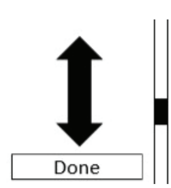
On the other hand, sharing an article from an untrusted source will cause your skill to decrease.



Liking a suspicious article or fact checking information from a legitimate source will get you few points. You may also use a hint, but it will cost you a few points.

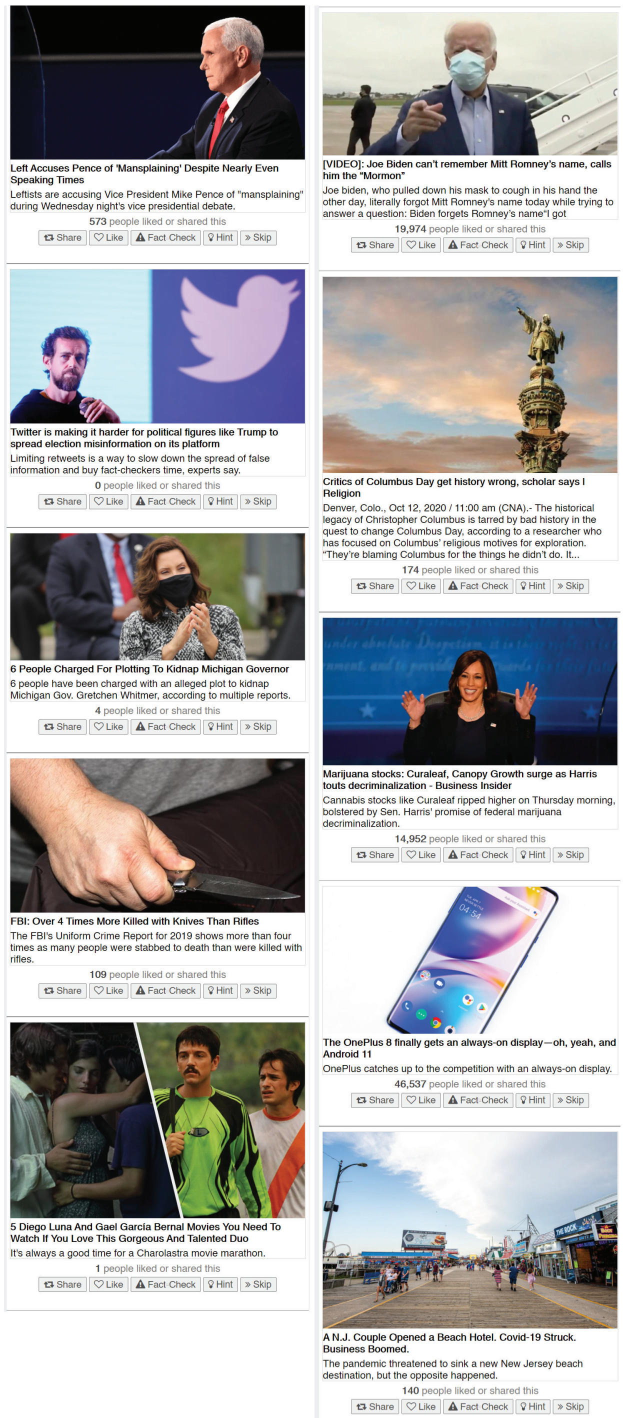
A.4 Page 4

Scroll down to view more articles. Scrolling past an article or selecting `Skip` will not affect your score.



To end the game session, scroll down to the bottom and press `Done` to see how you did. You may play as many rounds as you like. Check the leaderboard to see how your skills compare to other players!

B SAMPLE ARTICLES



C INTERVIEW QUESTIONS

[Show and explain the Study Information Sheet.]

Do you have any questions before we start?

May I record this session? *[If participant agrees, start recording.]*

As we discussed, the purpose of our research is to study a game called Fakey by gathering feedback from people such as yourself. Please feel free to be candid. Your honest feedback will help us understand how to make games like Fakey more effective.

[Bring up Fakey and share screen.]

As you can see, Fakey displays a variety of articles from mainstream as well as fringe sources. Your primary task will be to distinguish real information from questionable content.

Looking at each article, you are able to see the photo connected to the article, the headline of the article, a brief description of the article, and the number of users who liked or shared the content.

User Interface:
For starters, we would like to gather your initial thoughts on Fakey.

- What are your first impressions upon seeing Fakey?
- Go through each UI element (e.g., button, caption, etc.) within Fakey and ask: What do you think this element does?
- What do you believe is the purpose of including the photo with each headline?
- What would you say is the purpose of including the number of interactions (i.e., likes and shares) for each article?
- How do you think Fakey operates?
- What do you think is the purpose of Fakey?

Now that you've had a chance to think about Fakey, we would like you to experience interacting with it. Before we let you play with Fakey on your own, we would like to give you a brief overview of how each button within Fakey works and how it will affect the score you achieve:

- `Share` - The share button means that you would like to share this article with your social network. If you share accurate information, your Points and Skill will increase much more than if you share misinformation.
- `Like` - The like button indicates that you would 'Like' the article on social media. Liking a true article will increase your Points and Skill, though not as much as sharing it.
- `Fact Check` - This button is used if you believe an article may contain misinformation. Fact checking potential misinformation will increase your Points and Skill much more than doing so for articles from mainstream sources.
- `Hint` - If you press the `Hint` button, Fakey will reveal the source of the article. You can then use this extra information to help you distinguish between real or questionable news. However, you will get fewer points for taking the correct action with a hint than without a hint.
- `Skip` - As the name implies, the `Skip` button simply skips the article. You are free to skip any article.

Before we proceed, do you have any questions about how Fakey works?

Now, we would like you to interact with Fakey on your own. Fakey can be accessed by going to: https://fakey.iuni.iu.edu/. Please visit the site using any device and browser of your choice. When you are on the site, you may select 'Play Anonymously' if you do not wish to login to your social media account. We will not observe you while you play, so please take your time and play as much as you wish. We will wait while you do so. Please let us know when you are done playing.

[Wait while the participant interacts with Fakey.]

Now that you have played the game, we would like to ask you questions about your experience.

User Experience:

- What was your experience in interacting with Fakey?
- How did you decide whether to play anonymously or use your social media account? Why did you prefer one over the other?
- How well did the buttons in the UI match what you expected them to do?
- When did you use the `Hint` button? If participant did not use the button: why not?
- What do you think of the `Fact Check` button?
- How would you use the `Fact Check` button if it were available on mainstream social media? How often?
- What factors did you take into account when deciding how to interact with the articles presented in Fakey?
- What do you think about the Points and Skill you received from Fakey?
- How well do you think the Points and Skill match your ability to spot misinformation?
- What did you like about Fakey?
- What did you dislike about Fakey?
- What, if anything, was confusing or difficult during your interaction with Fakey? If anything was confusing or difficult: what would have made it easier?
- What do you think is missing in Fakey?
- If you could change anything about how Fakey works, what would it be? Why?
- If you could improve Fakey by adding any functionality, what would you add? Why?

Participant Background:
Finally, we would like to know a bit more about you and your general thoughts about online misinformation.

- What do you think about fake news and online misinformation?
- What are the major sources from which you get your news?
- How do you determine whether the information contained within a news article is accurate and trustworthy?
- What are your thoughts on fact-checking sites and services?
- Have you ever used fact-checking sites and services? If yes, when and why? If no, why not?
- Which social media services do you use on a regular basis? If participant does not use social media on a regular basis: why not?
- How often do you like or share news articles on social media? Why kind of news articles? With whom? For what purposes?
- How does the number of social media likes, shares, or comments influence your opinion regarding a news article?

Conclusion:
Those are all the questions we have for you. Before we end, is there anything else you'd like to tell us?

Thank you for participating in our study. We greatly appreciate your time. We will send you a $5 Amazon gift certificate to your email address.

If you have any questions after we end, please feel free to contact us by email. Thank you.

Received June 2020; revised October 2020; accepted December 2020