Classroom Activity for Critical Analysis of News Propagation Online

Christina Pollalis
cpollali@wellesley.edu

Catherine Grevet
cgrevet@wellesley.edu

Lauren Westendorf
lwestend@wellesley.edu

Samantha T. Finn
sfinn@wellesley.edu

Orit Shaer
oshaer@wellesley.edu

See next page for additional authors

Follow this and additional works at: https://repository.wellesley.edu/scholarship

Version: Publisher's version

Recommended Citation
Pollalis, Christina; Grevet, Catherine; Westendorf, Lauren; Finn, Samantha T.; Shaer, Orit; and Metaxas, Panagiotis, "Classroom Activity for Critical Analysis of News Propagation Online" (2018). Faculty Research and Scholarship. 189.
https://repository.wellesley.edu/scholarship/189

This Conference Proceeding is brought to you for free and open access by Wellesley College Digital Scholarship and Archive. It has been accepted for inclusion in Faculty Research and Scholarship by an authorized administrator of Wellesley College Digital Scholarship and Archive. For more information, please contact ir@wellesley.edu.
Classroom Activity for Critical Analysis of News Propagation Online

Abstract
We present an educational activity for college students to think critically about the truthfulness of news propagated in social media. This activity utilizes TwitterTrails, a visual tool to analyze Twitter claims, events, and memes. This tool provides views such as a propagation graph of a story’s bursting activity, and the co-ReTweeted network of the more prominent members of the audience. Using a response and reflection form, students are guided through these different facets of a story. The classroom activity was iteratively designed over the course of three semesters. Here, we present the learning outcomes from our final semester’s evaluation with 43 students. Our findings demonstrate that the activity provided students with both the conceptual tools and motivation to investigate the reliability of stories in social media. Our contribution also includes access to the tool and materials to conduct this activity. We hope that other educators will further improve and run this activity with their own students.

Author Keywords
TwitterTrails; fake news; social media; educational activity; misinformation.
**Introduction**

The term “fake news” became increasingly popular in the media recently, though the problem of online misinformation is not a new phenomenon [16, 17]. The spread and accessibility of social media has allowed every online account (either human or bot) to circulate any claim, including fabricated claims masqueraded as news, without much verification or accountability. As such, false information can be easily perpetuated, leading to a spread of misinformation [18]. Several studies indicate that social media users do not currently have the appropriate tools to critically assess whether a claim is true, and that the spread of misinformation can lead to a belief of fake news as truth. For example, a Stanford University study [20] found that 82% of middle-schoolers were unable to distinguish between a “sponsored content” ad and a real news story. Findings from this study further indicate that rather than judging a social media post based on source credibility, students judged it based on the amount of detail the post contained or the size of the photo attached [20]. For example, approximately 40% of the high-schoolers in the study believed that a deformed daisies photo posted on the photo-sharing website Imgur with the title “Fukushima Nuclear Flowers” and a subtitle of “Not much more to say, this is what happens when flowers get nuclear birth defects” meant that the area near the Fukushima Daiichi nuclear plant in Japan was toxic, despite a lack of the post source or any news story attached [20]. This inability to distinguish real and fake news is especially problematic for young adults, since studies indicate that a majority of young adults over the age of 18 receive their news from online social media, such as Facebook or Twitter [12].

Unfortunately, this issue goes beyond gullibility and can have real-world outcomes, as it has been argued that the propagation of fake news may have played a vital role in the 2016 U.S. Presidential Election [2, 23].

With the simultaneous rising of “fake news” and the increasing inability to determine the validity of a particular social media news post, it is becoming vitally important to train young adults to distinguish between real and fake news. There have been multiple attempts to facilitate this distinction, both by social media platforms, news outlets, and game inventors. Facebook released a tool that flags disputed news stories and warns a user before they share such a story [4]. Google released Fact Check, a tool implemented in their search and news, that summarizes the article’s claim, who it was claimed by, and whether a reputable source labeled the claim as true or false [6]. Certain news outlets have released truth verification tools, such as BBC’s Reality Check [5] and The Washington Post’s fact checker [26]. There are also third party sites that act as fact checkers, such as Snopes [24], PolitiFact [19], and FactCheck.org [8]. Furthermore, there are domain-specific fact checkers, such as SciCheck [9], which fact checks science-based claims, Health Watch [7], which fact checks the health-care debate, and Trump Transcripts [10], which fact checks President Trump’s remarks. Lastly, American University Game Lab/JOLT has created a game that helps players sort fake from real news [3].

In this case study, we present an in-class educational activity for college students (see Figure 1 for a flow chart of the activity). The activity uses TwitterTrails...
[22, 28], a novel online visual tool for investigating the trustworthiness of stories (claims, events, and memes) spread on Twitter. TwitterTrails analyzes and visualizes information from tweets related to a specific story. Through the presentation of various interactive visualizations, including a Propagation graph (Figure 3), a Time Series of Relevant Tweets (Figure 2), and a co-ReTweeted network (Figure 4), TwitterTrails provides users with means and evidence for evaluating the trustworthiness of a claim.

The goal of the educational activity we present here goes beyond the use of TwitterTrails for assessing a particular claim. We demonstrate how to provide students with a conceptual toolbox for evidence-based inquiry of the reliability of online information.

TwitterTrails
TwitterTrails (http://twittertrails.com/) [22, 28] is a novel online visual tool, developed by some of the co-authors, that allows users to investigate the trustworthiness of a “story” (a claim, a meme, or an event) shared on Twitter. Prompted by a search with relevant keywords using the Twitter Search API, TwitterTrails (mentioned as “TT” from now on) collects and analyzes all tweets matching the given keywords automatically. While it does not answer directly the question of a story’s validity, it provides detailed evidence for examining how a story propagates on Twitter and how Twitter users participate in spreading the story. Specifically, the tool includes algorithms measuring the spread of the story, the terms used to self-describe groups of influential spreading groups. TT automatically produces a page for each story investigated based on the data calculated by its algorithms, which indicates: the investigation date; the first relevant tweet and its author; when the story “broke” on Twitter, that is, when it received significant attention; the amount of time it took for the first “breaking tweets” to be posted; the total amount of (verbatim) retweets; tweet variations that could be posted by spammers; and the incoming tweet rate of the story spreading. The tool also displays interactive visualizations: of the Propagation Graph (story’s bursting activity) (see Figure 3); of the Time Series of Relevant Tweets (see Figure 2); of the co-ReTweeted network [21] (see Figure 4); and relevant statistics including word clouds of the most common words appearing in the profiles of influential spreading groups.

TT is implemented in Python using the Twitter Search API for data collection, the Gephi Toolkit [11] for data analysis, and a MongoDB database. The story page and visualizations are produced using PHP and JavaScript with libraries such as Highcharts [12] for the time series and propagation graphs, and Sigma.js [24] for the co-Retweeted network. At the time of this writing twittertrails.com hosts 550 stories, each containing up to 200K collected tweets per story, and is hosted on Amazon AWS.

The current version of TT also has two metrics for each story: spread, which measures the visibility of a story; and skepticism, which measures the ratio of tweets containing negative terms over the overall tweet volume. These measures were not used in this study. They will be part of a future study evaluating the effectiveness of Machine Learning algorithms to detect the validity of a story. We did not use these metrics or the algorithms in the activity we describe below.
The propagation graph visualization presents a close up at the time when the rumor became more visible, which happens by attracting tweets with a relatively high number of retweets. It is composed of bubbles (tweets) spread over a period of time that TT deems to represent the earliest bursting of the story. Each bubble is a tweet, the higher it appears on the graph, the more retweets it attracted. The diameter of a bubble represents the tweet potential, computed as the log of the number of the tweeter’s followers. In this story, the tweet with the highest burst is the one posted on July 7, 2015 by a tweeter (Dinesh D’Souza) who had 200K followers. That tweet, which happened to be the same tweet that initiated the TT investigation, had 1638 retweets and 1274 favorites. (Interestingly, all the earlier tweets in the propagation graph shown above are claiming that the photo is fake, yet they were ignored by D’Souza!)

**Use scenario**

As an example of a TT use scenario, consider the following story, which is one of 12 stories we used in our in class educational activity. The story investigates a claim that “Hillary Clinton is pictured with the confederate flag.” The basic elements of the story are shown in Figures 2-7.

**In-Class Activity**

We developed a new educational activity that utilizes TwitterTrails for investigating the reliability of stories on social media. Here we share the activity and lessons learned from running the activity in four sections of an introductory Media Arts and Sciences course at a liberal arts college, spanning over three semesters. Students taking this class do not have a technical background, as it is an introductory course for non-Computer Science majors. In all three semesters, the students had prior discussions about the Internet and social media. This activity was presented as an in-class exercise to explore these concepts in a real context.

**Learning Goals**

We defined the following learning goals for our in-class activity. In particular, we expect that following participation in the activity, students will be able to:
Understand the concepts of rumor spreading, including the extent and mechanisms in which stories propagate over social media.

Read and interpret visualizations that describe propagation of information over time.

Conduct an evidence-based inquiry into the reliability of online information, employing a set of questions to examine who is spreading a story, and when and how was the story propagated.

Identify indicators and characteristics that impact reliability including polarization, echo chamber, timing, and Twitter bots.

Design Process
We designed the activity using an iterative process. The two sections from the first two semesters, with 19 and 10 students respectively, served as a formative evaluation, informing our design of the tool and activity. In these semesters, we grouped students into teams of 3-4 students, so that each team would use TT to investigate two pre-curated stories each. We dedicated about 40 minutes to the task, following a 30-minutes training. Each team was assigned a relatively simple celebrity death story, followed by a complex story involving a rumor about a recent event or a political figure. Each team was required to fill out a questionnaire (implemented as a Google Form) for each story. The questionnaire guided students in the process of investigating and identifying evidence regarding the truthfulness of the story. Following the activity, each student was required to submit a reflection questionnaire.

Based on our findings, we simplified TT’s interface, removing some of the functionality for filtering and zooming on the propagation graph and time series, so that the important information is clearly presented and no extensive user training is needed. Moreover, we removed the results of the Machine Learning algorithm, which calculates whether the rumor was true or false. In the third run that we are describing below, we also used new stories, switching from celebrity death rumors to more challenging ones.

The Twitter Trails activity
Following the redesign, the activity consists of an 11-minute training video [15], and an activity form, which guides students through the investigation of a story using TT. Groups are assigned two stories each, and are given the task of investigating the origin, spread, and validity of each story. The activity form guides the investigation with a series of open-ended questions about a story’s propagation graph, the time series of relevant tweets, the story’s co-retweeted network, and the story’s most retweeted images. Lastly, students are asked to estimate how valid the story is, and to describe how they reached that conclusion. This process can be found as a flow chart in Figure 1.

The activity is intended to last about 25 minutes per story so that students investigate at least two different stories in class (total of about 50 minutes of active investigation). Following the in-class investigation, each student is asked to spend 10-20 minutes answering an individual reflection form. The reflection form consists of open-response questions about what students learned from the exercise and how what they learned can be applied to their life and future career.
The stories we integrated in the activity are listed in Table 1. We selected these particular stories based on their recent timeliness, the clarity of claim, and the ambiguity of validity. We included stories that are timely and relevant so that students experience the tool in a real-world setting. In addition, students may not have formed a strong opinion yet on the veracity of recent stories. We chose stories that have clear claims relating to a specific fact, and have a definitive validity. This allows us to evaluate students’ learning outcomes following the activity. We also selected stories that are relatively not well-known through media coverage and thus students might not have an a priori answer regarding their veracity. Finally, we chose stories that represent multiple sides of the US political spectrum. We also balanced True and False claims, and recommend that each team investigates at least one True and one False story.

It is important to note that instructors can curate new and timely stories for this activity using TT.

The training video, in-class activity form with curated stories, and the reflection form are all available online here: http://bit.ly/2kDpt66

Instructions for creating new stories using TT are available here: http://bit.ly/TTrequest

**Evaluation**

In the third semester (Fall 2017), we ran the activity with 43 first-year and sophomores in two separate sections, with 22 and 21 students, respectively. Before the activity, students were presented with an 11-minute in-class introduction on TT. In both sections, students broke into teams of 3-4 students and were assigned a group number from 1 to 6. Each group was assigned two stories from Table 1 (one true and one false) based on their group number, and were asked to follow the activity form for each story. In this form, each student was asked to answer 16 comprehension questions, identify whether the claim was true or false, and provide the pieces of evidence that support their evaluation. The total time allotted for the task was 20 minutes per story.

**Title of story**

- Rush Limbaugh evacuates Florida after claiming that Irma is a liberal hoax
- A shark was spotted on a Houston freeway during hurricane
- Facebook sold political ads to fake Russian account
- Obama went golfing during Hurricane Katrina
- Candidate for Education Secretary DeVos wants guns in schools because grizzly bears exist
- Shiva Ayyadurai invented email
- Hillary Clinton is pictured with the confederate flag
- Harvard hired Chelsea Manning as a Fellow
- Black Lives Matter block Hurricane Harvey rescue efforts
- Trump makes Computer Science education a priority
- Posting a copyright status message on Facebook will protect your posts from copyright violations
- White House used private emails

**Table 1**: The list of 12 stories used for the evaluation. The titles link to TT story pages. The false stories are shown in red.

**Learning Outcomes**

All students completed the task in the allotted time, submitting 86 activity forms (each student submitted
forms for two stories). A large majority of students were able to accurately evaluate the claims during the activity. Out of the 86 responses, 88.3% of the students correctly assessed the validity of the claim. For those who did not make a correct assessment, 5.8% were not sure about the validity of the claim, and 5.8% incorrectly assessed the validity of the claim.

We used content analysis methods to analyze student open responses. We analyzed the evidence students provided to justify the evaluation of a claim. The most popular piece of evidence used was the tweeter’s identity, mentioned in 66.3% of responses mentioning it, followed by the content of the tweet, mentioned in 48.8% of responses, and the retweeted images, mentioned in 16.3% of responses. For this question, 41.9% of students provided 1 piece of evidence, 39.5% provided 2 pieces of evidence, 12.8% provided 3 pieces of evidence, and 2.3% provided 4 pieces of evidence. These findings indicate that students were able to conduct an evidence-based inquiry into the reliability of online information (L3).

Reflection
In response to the question of what they learned from this activity, 52.4% of the students mentioned learning how to assess the validity of a claim (L3, L4). In the voice of one student, “I learned that you can determine the validity of a claim made on twitter by following its history. Additionally, there’s a lot to make out of each tweet, such as when it was posted, who posted it, how many retweets, etc.” 45.2% of students mentioned learning about how false stories are propagated through social media. For example, one student wrote, “I learned that information--regardless of its level of truth--can spread incredibly fast across the internet.

And many of the sources that make a false claim can do so convincingly enough to gain a lot of traction in a short amount of time” (L1). 42.9% mentioned learning about using the TT visualization tool. One student wrote, “I learned how to use TwitterTrails to follow a story that is broken on Twitter. I learned how to detect whether a story is likely true or not based on the graphs and tools on the program” (L2). Lastly, 11.9% mentioned learning about the twitter platform and how it can perpetuate fake stories. As one student wrote, “Twitter can be an echo-chamber if the information/news [you’re] following does not "break" and get rightly supported or refuted by reputable sources. This can be especially dangerous if Twitter is a person’s main source of news in a polarized, narrow, and or isolated news-community” (L4).

97.7% of students reflected that this exercise is applicable to their lives and future careers. Specifically, 62.8% of students responded that the fact that they learned to assess claims online will help them in the future. In particular, one student wrote ”I think it will help me think critically, and think twice about what I see online. It can help me better analyze news that I find on Twitter.” 32.6% reflected that the activity made them understand the importance of assessment and that they will think critically about claims online in the future. For example, one student wrote “I am interested in public policy, so it is very helpful and relevant for me to understand how and why certain stories get spread and how to assess the validity of a story.” 23.3% felt that their understanding of fake news increased. One student wrote, “I use twitter everyday, and using this tool makes me more conscious of the things I connect my account to (whether that be by favorite, retweeting, or mentioning
content).” Finally, 21% reflected that they will be using TT to analyze online social media claims in the future. For example, one student stated, “If I am ever questioning a major news story, I can use TwitterTrails to assess its validity.”

Lessons Learned
Reflecting on the learning outcomes and on the activity itself, we identified several lessons. First, in future iterations, we will start the activity by asking the students to try to determine the validity of the claim just by reading the title of the story before showing the TT investigation. Through our evaluation, we found that a few students were misled by the framing of the title of the claims they were analyzing. By isolating their understanding of the title itself, we will be able to better determine the impact the TT tool has on the students’ ability to determine the validity of a story.

Second, the stories we used in this evaluation had a clear veracity. This allowed us to evaluate student learning using TT and measure our learning objectives. However, many stories, especially those that are evolving in real time, are not as clearly determined. A future goal is to elaborate this activity to include stories that do not have a definite correct or incorrect validity.

The tool that the most students (51.2%) found helpful was the propagation graph, followed by the co-retweeted network and statistics (39.5%) and the time-series graph (34.8%). This indicates that no singular view is clearly the most helpful, and that it is important to provide multiple tools and perspectives so that different users could successfully investigate diverse stories. We also learned that while the co-retweeted network deemed helpful, it was the least easily understood among the views (average Likert score of 3.26 for “easy to understand,” where 1 is “Strongly Disagree” and 5 is “Strongly Agree”, SD=1). Future iterations of the tool and activity, will improve the usability and understandability of this view.

Finally, an important takeaway is that young adults, when provided with appropriate tools, are able to learn a process for evaluating the reliability of stories in social media. Our instance of the activity relied on Twitter data since 1) Twitter is often the source of breaking news, and 2) the Twitter API supports a tool like TT. However, the learning outcomes are not specific to Twitter as students reflected on the propagation of news in social media more broadly. This activity structure could also be adapted for other social media platforms, provided that an evidence-based analysis tool can be developed for them. We found that by empowering students with both tools and a process to investigate social media stories, this activity can motivate students to apply a similar process and to seek these types of tools in their lives and future career.

Conclusion
We presented an educational activity for college students that utilizes TwitterTrails, a novel online visual tool to teach students how to investigate the reliability of stories on social media. Our findings demonstrated that the activity provided students with both the conceptual tools and motivation to investigate the reliability of stories on social media. We hope that other educators will further improve and use this activity with their own students.
References


