The Absence of a Media Literacy Toolbox: Working toward an Evaluation Tool.

Gabriel Yakub, Simon Knight, Kirsty Kitto, Peter Fray

Word Count: 6968

Key Words: Media Literacy; Technology; Critical Thinking; Fallacy; Strategies;

Abstract

This paper discusses the fragmented nature of media literacy and its relationship with technology. It highlights the need for standardised media literacy strategies, particularly the strand that deal with evaluation, which can help address the challenges of the current media landscape, for instance, the fake news phenomenon. Subsequently, we introduce early work towards developing a new evaluative media literacy tool that can empower media consumers to think strategically about the information they are exposed to. This tool, called Fallasigns, is based primarily on research that suggests news topics can attract specific logical flaws. Fallasigns cultivates the ability to anticipate the most likely logical and rhetorical pitfalls to emerge in a news story before being exposed to it. We argue that this strategy may work to effectively inoculate media consumers and provide a more systematic approach in evaluating information.

Introduction

As information is easier to access, produce and distribute, consumers are now, more than ever, required to make faster decisions about the quality of the increasingly large quantity of information they are exposed to online. This has implications for how current media literacy strategies can remain effective, and be utilised by media consumers and educators (a point to which we will return shortly). Therefore, this paper has two objectives. The first is a discussion of the fragmented nature of media literacy and its relationship with technology. This fragmentation highlights the need for standardised media literacy strategies, particularly those dealing with the evaluation of information, which can help address new challenges that have arisen in the media landscape, such as the fake news phenomenon. These strategies are important for media consumers, and educators, because although the specific content of media changes there may be common media literacy strategies to tackle these issues. The second objective is to introduce a proposed media literacy tool that can empower media consumers to think strategically about the information they are exposed to. This proposed tool, called Fallasigns, is based on a premise, supported by preliminary research, that news topics can attract specific

logical flaws. Supporting media consumers to identify or anticipate these fallacies would provide them with a more systematic approach to evaluating information, potentially helping to inoculate them against cognitive fallacies in media. In this paper we start by providing an overview of the conceptualisation of media literacy and prior work in this space. We note that work has been fragmented in nature, and faces challenges in the contemporary media environment. We then describe empirical work which outlines an approach to supporting media literacy through the proposed tool, and the grounding for that tool.

Media Literacy

Media literacy is multi-faceted and interdisciplinary, and can mean many things to many people, as such it is often difficult to come to a single agreed upon definition (Brown, 1998; Koltay, 2011). A proliferation in use of the term occurred in the 1990's which led to the emergence of many diverging conceptualisations (Brown, 1998; Buckingham, 2013; Hobbs, 1999; Hobbs & Frost, 2003; Hobbs & Jensen, 2009) and since then the term has been hotly debated. This debate can be attributed to three underlying problems. Firstly, 'literacy' can be understood in two ways. The word 'literate' is typically ascribed to one who can read and write, and so media literacy can be understood in terms of basic comprehension and production skills, or 'functional literacy' as it is sometimes referred to. However, 'literacy' can also be understood as 'critical literacy', having a "broader analytic understanding" of "different modes of communication" (Buckingham, 2013, pp. 37, 38; Hobbs & Frost, 2003). The underlying definition of literacy that is adopted can lead to markedly different expectations about what 'media literacy' should imply. Secondly, the understanding of media literacy changes with the media environment, which is to say, as technology changes new competencies must be developed to use them. People must acquire an understanding of "the unique grammar of each medium", i.e. the way a medium creates meaning with its unique way of representing information (Meyrowitz, 1998, p. 103). For example, jump-cuts, soundtrack and colours can be used to shape meaning in a film. By contrast, words are symbols that have an abstract relationship to the ideas they signify, where meaning is constructed through the relative positioning of words to each other, together with the use of punctuation. The third and final problem arises when we consider that media literacy is sometimes used as an umbrella term that encompasses "various subsets... such as information literacy, digital literacy, critical literacy and news literacy" (Ashley, Maksl, & Craft, 2013, p. 8). Those who adopt this form of definition will have markedly different expectations about what media literacy implies when compared to some of the more restricted definitions discussed above.

Druick (2016) claims that typically, media literacy is understood as analytical, and specifically the "critical study of the media" (p.1126) or the application of critical thinking, and not just a functional skill (i.e. reading text, writing, using a computer etc.). This is because texts are not always straightforward messages. Embedded within them are ideological structures that can attempt to persuade, mislead or shape one's perception of reality, in obscure ways (Fiske, 2010).

Therefore, media literacy can be examined within the scope of the political, economic, social, cognitive and technological (Bawden, 2001; Livingstone & Helsper, 2010; Potter, 2016).

The role of technology in the development and understanding of media literacy is particularly pertinent, given the drastic shift of the information environment within the last two decades, which has seen the unprecedented ease of access, production and distribution of information (Buckingham, 2013; Koc & Barut, 2016). These changes have brought with them challenges to media consumers, requiring a re-examination of the efficacy of traditional media literacy strategies. As such, we will discuss below the problems of access and production within the current media environment, emphasising how recent technological developments in the media landscape have complicated the demands of media literacy.

Access and 'Prosumption'

The current media landscape can be defined by major shifts in access, production and consumption. Prosumption, the portmanteau of 'production' and 'consumption', describes this interrelationship between what is produced in a society based on what is consumed (Islas, Arribas, & Gutiérrez, 2018). Today, the media landscape reflects this symbiotic relationship of individual production and consumption (Garcia-Ruiz, Ramirez-Garcia, & Rodriguez-Rosell, 2014), with web 2.0 technologies affording the production and dissemination of user generated content that facilitates this prosumption (Ritzer & Jurgenson, 2010), with broad access to smart devices affecting how society interacts with information (Lee, 2016). Moreover, a landscape where social influencers widely express and relate ideas and values in real-time, can be consumed and re-produced or 'shared' instantly. This calls for a media literacy that can address the repercussions of 'sharing' ideas and values that can have a wider community impact, as an MIT study has found that "false news spreads further, faster and deeper, and more broadly than truth in all categories of information" (Vosoughi, Roy, & Aral, 2018, p. 2). Consequently, "information and strategic skills" are required to navigate information available online (Van Dijk & Van Deursen, 2014, p. 57).

The technology that makes up this developing and shifting communication landscape is also often referred to as new media, which, aside from its participatory nature, includes the production of information from innumerable online users, as well the construction of multifarious online spaces that represent many interests and ideas not typically or widely represented in traditional media (Cope and Kalantzis 2010). In an effort to understand new media and subsequently what would be considered as new media literacy, Chen, Wu and Wang (2011) constructed a theoretical framework, later developed by Lin, Li, Deng and Lee (2013) that further characterises media literacy in terms of both consumption and prosumption, taking into account both the critical/analytical and functional skills required to efficiently operate in this new environment. Given these theoretical developments, inn what follows we seek to locate and

discuss the specific standardised media literacy tools available to educators and online users alike.

Media Literacy: An Imaginary Toolbox?

There are many curricula and suggested methods that exist under the banner of media literacy. However, we are yet to see a standardised approach or toolbox that educators can decisively draw upon to cultivate media literacy skills in any given context (Mujica, 2012). Indeed, many of the tools educators have traditionally relied upon are fast becoming irrelevant as our media landscape becomes more complex. For example, media literacy rubrics like the CRAAP test (Currency, Relevance, Authority, Accuracy and Purpose), have become outdated in the current sophisticated media landscape (see the expectancy violation heuristic below). One of the reasons why a standardised media literacy toolbox does not exist lies in the problem identified by Brown (1998) who argues that media literacy varies with pedagogical context. In other words, what is taught is relative to the demographic and institutional context of the education, which for instance could include a, "school district, university degree program or regional government agency" (Brown, 1998, p. 48). Another reason lies in the lack of an agreed upon understanding of media literacy, as well as the contextual relationship that exists between audiences and the media. This is because individuals and social groups utilise media "to meet their own interests and needs" (Bernardi, 2016, p. 25). Despite these varied approaches, among scholars there are widely accepted definitions and basic principles that provide a platform from which educational methods can be produced. For example, the definition Aufderheide (1993) proposed, has come to be widely accepted:

"A media literate person - and everyone should have the opportunity to become one - can decode, evaluate, analyse and produce both print and electronic media. The fundamental objective of media literacy is critical autonomy in relationship to all media. Emphasis in media literacy training range widely, including informed citizenship, aesthetic appreciation and expression, social advocacy, self-esteem, and consumer competence" (p. 9).

Marten (2010), however, states that this definition lacks sufficient detail for media educators to "design strategies" around, citing an approach earlier championed by Potter, who emphasised "knowledge structures", that is to say a psychological process that promotes the organisation of information in meaningful ways. On the other hand, Kellner and Share (2005) emphasise a more epistemological approach, drawing on standpoint and feminist theory, which sees dominant values and perspectives in society de-naturalised, an approach understood as critical media literacy (Kellner & Share, 2007; Ávila & Pandya, 2013).

Despite discrepancies in emphasis and definitions among scholars, overall there is a common understanding that media literacy, at a fundamental level, is the application of critical thinking to media texts. Even though Potter (2016) argues that the term 'critical thinking' is ambiguous, he outlines its defining features in order to clarify what critical thinking actually means: synthesis,

analysis, evaluation and abstraction (p. 17). Thus, the point can be made that media literacy, unlike the branch of logic in philosophy, does not provide a widely shared or agreed upon prescribed set of analytical tools or formulas, but rather describes the function and outcome of what media literacy should look like. However, this can be problematic for educators who attempt to create appropriate tools for a given context using their available (and often limited) resources.

Therefore, aside from the flux of technological change, this state of affairs heavily accounts for the fragmented nature of media literacy, which sees the existence and production of many strategies and perspectives. This is present in both the Australian and the UK educational contexts. In Australia, the study produced by Nettlefold and Williams (2018) recommends that, "Future researchers, policy makers, and educators need to focus more on standards of teacher training and professional development in media education" (p. 10). More specifically, in the UK, the National Literacy Trust has shown that 'stakeholders' in media literacy education are calling for "shared definition of critical literacy and clear learning outcomes". The reasoning provided here is that, "evolving definitions of critical literacy provide a context within which critical literacy skills in the digital age may be better understood" (National Literacy Trust, 2018, p. 17).

Media Literacy and the Complexity of Information

In bringing together the many strands of media literacy, in their 'Field Guide to Media Literacy in the United States', RobbGrieco and Hobbs (2013) place critical thinking as the fundamental link between them all. Much like Potter, here, critical thinking is understood as the ability to access, synthesise, analyse and evaluate information (Ashley et al., 2013; Buckingham, 2013; Fleming, 2014; Potter, 2016). The goal of media literacy educators and scholars has recently focused on improving the skills of media consumers to identify false information, and specifically, the ability to evaluate information for its validity (Mihailidis & Viotty, 2017; Nettlefold & Williams, 2018; The News Literacy Project, 2019). This is because the current media environment is characterised by the "fake news phenomena" (Bakir & McStay, 2018; Burshtein, 2017, p. 398). It then begs the question - just how effective is media literacy under these circumstances?

Does Media Literacy Work?

Some researchers argue that media literacy contains embedded assumptions that make it ineffective from the outset. For example, in relation to preventing consumers from sharing fake news, Marwick (2018) argues that media literacy is not effective against those who already have a strong distrust against the media, noting instead that the affective or emotional appeal of a story often overwhelms any consideration of its factual accuracy, stating, "in many cases, what matters is the affective or emotional appeal of a particular story or claim, rather than its factual accuracy." (p. 509). This is particularly problematic, because social media, where news consumers are more likely to find their news (Park, Fisher, Fuller, & Lee, 2018), "allow[s] for

the rapid amplification of emotionally-charged messages across platforms" (Albright, 2017, p. 88). Furthermore, one of the distinguishing features of fake news is in fact its ability to provoke an emotional response in the reader as well as to hold attention and gain more viewership (Bakir & McStay, 2018). In response to this proliferation of disinformation, media producers have attempted to use fact-checking tools to set the record straight, however, in many cases this has produced a reverse effect that has resulted in polarisation and selective exposure (Shin & Thorson, 2017).

As such, some media literacy organisations around the world have developed curricula that would help address these contemporary issues facing media consumers. In Ukraine, a new curriculum was created through the global not-for profit organisation Irex, called Learn to Discern (L2D). L2D provides training to local communities on how to discern disinformation from reliable information within the current media landscape in response to a perceived inability in current media literacy curricula to address the types of challenges we face today with disinformation and propaganda (IREXdc, 2017).

Critical Thinking: Evaluation

The dominant approach being adopted in developing media literacy curricula aims to improve evaluation skills in the hope that this will help media consumers to determine what information is reliable and what is not. This is not only because some media producers get things wrong, but it is also an attempt to combat the growing tendency for some producers to systematically produce information that is false, while making those online artefacts appear trustworthy. As a result, educators and scholars are concerned about the way media consumers determine what is credible online and how signs of trustworthiness or credibility cues are exploited. In their study Metzger, Flanagin, and Medders (2010) note that the expectancy violation heuristic is quite prominent among online users. This heuristic refers to a cognitive bias that individuals perceive online content or websites as not credible if expectations are not met in terms of content and appearance. This may include the layout, font and potential claims or ideological slant of the information presented on the website. This is one way that credibility cues are exploited, as fake news propagators typically try to produce online sources that pass as reliable and trustworthy.

Studies have found that relying on these types of checklists (which may also include: the absence of typos, a reference list, good web design, easy to find in popular search engines, etc.) is an almost sure way to get duped. In one study, college students and historians were asked to determine the credibility of certain websites. Their approach was compared to professional fact checkers, who were able to evaluate the same websites with speed and accuracy better than their counterparts in the study. From this result, Wineburg & McGrew (2017) concluded that the best way to avoid major pitfalls is to "evaluate unfamiliar sites by leaving them" and "finding out what the rest of the web has to say" (p. 45). Breakstone, McGrew, Smith, Ortega, and Wineburg (2018) added to this study, arguing that given "the health of our democracy depends on access to

reliable information", and the sophistication of misleading websites are near indistinguishable from credible content on the web, the ability to evaluate information is of utmost importance, and verification of information online is best done by 'reading laterally' (p. 31).

Evidently, the obstacles which the current media environment presents to educators and media consumers are challenging. In addition to this, Nettlefold & Williams (2018) have found that media literacy education is not taught consistently within the Australian Curriculum, with 50% of teachers surveyed claiming they only sometimes explore critical engagement in the classroom, 24% say they rarely do, 7% say they never do, while 19% say they often do so, although the authors suggest that this may be in fact due to limited time constraints within the classroom. Similarly, research in the UK reveals that 50% of teachers advise they deliberately teach critical literacy, however, only 26% say they do so regularly, while 40% of teachers say they do not teach critical literacy. Much like Australia, one of the most pressing factors is time constraint, as teachers require training and additional resources to provide effective learning opportunities to their students (National Literacy Trust, 2018, pp. 21-24).

Working Toward a New Media Literacy Tool

The discussion above suggests that there is a need to transcend the fragmented approach to media literacy. We require a new approach, one that addresses the limitations faced by educators in the classroom context, and in dealing with the fake news phenomenon, as well as the novel challenges that emerge with new media. We now propose a media literacy tool that supports critical thinking skills in evaluating the validity of information in an effective and timely way (Potter 2016). Our tool aims to support individuals in identifying the signature fallacies of a news story, which we believe would help people to guard themselves against potential logical pitfalls or rhetorical techniques aiming to swing them towards believing false claims. For clarity, a signature fallacy is the recurring, or likely occurring fallacy associated with a topic or a news story. We claim that this approach has potential to inoculate media consumers, as inoculation theory states that, "individuals can be inoculated against persuasive attacks on their attitudes in a similar manner to the way individuals can be immunised against a virus" (Banas & Rains, 2010, p. 283). Therefore, this tool may help to strategically deal with the fast-paced nature and high volume of information available in the current media landscape. As such, there are two ways our proposed tool could be employed, both online and offline.

The Online Approach

The online approach would require the production of an online tool, such as a web-browser plug-in. In effect it would run in the background and identify logical fallacies in a particular news topic the user is interested in, and relay to the user the signature fallacies they should watch out for in a particular topic (such as headlines that include Climate Change), perhaps using annotations or markup on a webpage, or perhaps in a pop-up window. If the online user is unfamiliar with the different kinds of fallacies they might be exposed to, then the program could

provide simple explanations, however, this tool is intended to work as a kind of digital heuristic, whereby the program would use machine learning to identify and collate the logical flaws that appear within a topic area, and display the recurring fallacies related to that topic, for the online user. This would effectively allow a user to make decisions about the reliability and quality of available information speedily.

The Offline Approach

In order to determine the likely signature fallacy of a topic, one must exercise their ability to abstract, analyse and evaluate information, since one needs to be aware of the basic components of a news story and their implications in order to anticipate potential flawed arguments. This can be done in a number of ways, as a: simple thought experiment, turned into a mind-map exercise, or discussed in a class or group setting. A basic example of this approach can be demonstrated in just a few steps:

1. Select a topic

2. What specifically would people take sides about, in this topic, or, what are the main areas of contention?

3. What logical flaws would people likely fall into either defending or addressing the opposite side of the issue? For example, if the topic at hand is scientific in nature, the participant may well expect questions around cause and effect, who are key thinkers, or what data is used to support the claims in the discussion.

For example:

1. Topic: Climate Change

2. Main contention: Are humans causing climate change?

3. Possible logical fallacies: appeals to authority, cherry picking, false cause...

The strategic thinking promoted in the offline approach would seem to enhance the effect of inoculation. In the context of news consumption, a study by Cook, Lewandowsky, and Ecker (2017) found that when participants were informed of the techniques used to produce misinformation prior to their exposure to it, its polarising effects were neutralised, suggesting that “inoculation interventions boost strategic monitoring when encoding potential misinformation”, increasing the depth of thinking from a heuristic approach of evaluation (p. 15, 21). As such, this offline approach may prove to be more effective with respect to inducing inoculation against misleading or erroneous information, as it asks the participant to implicitly

understand what arguments are and how fallacies work to produce unreliable conclusions. Moreover, unlike the online approach which would aggregate and display fallacy signatures, the offline approach asks participants to anticipate what these are - promoting greater strategic thinking with respect to the nature of the issue at hand. As such, it is predicted that the offline approach may be useful, not just for news consumers but also educators, as this method can be turned into a class discussion or a solo brain-storming activity for students. For further clarity around the importance of understanding what logical fallacies are and how they work, a brief overview is provided below.

Logical Fallacies

A logical fallacy is a flaw in reasoning where a conclusion does not follow from the underlying premise/s proposed to support it with any certitude or likelihood. Such fallacies can be understood as an argument that appears to be valid, but in fact is not (Hansen, 2010). Fallacies are classed into two categories, formal and informal. Formal fallacies pertain to logical flaws committed in deductive arguments, such as syllogistic or propositional arguments (Vleet, 2011). They arise when an argument is rendered invalid by a flaw in its logical structure, that is, the underlying semantics of the statement need not even be considered. These types of arguments are only valid if there is no way the conclusion can be false if the premises are true. This solely depends on the relationship between the premises, which are determined by their location in the argument. Therefore, validity is judged on the basis of structure. Take for example the propositional argument If A then B, A therefore B. As a popular example, this could read, 'If it rains then the ground is wet, it rained therefore the ground is wet.' It need not have rained for this argument to be valid. The invalid form of this argument would read like this 'If it rains then the ground is wet, the ground is wet, therefore it rained.' This latter form is invalid because it is now phrased in such a way where the consequent i.e. the 'wet ground' could have been caused by any number of things, however the former proposition implies that the wet ground must have been caused by rain.

An informal fallacy relates to inductive arguments, the fallacies that emerge typically have less to do with the structure and more to do with the content of a statement and its correspondence to reality (Vleet, 2011, p. 8). 8). Take for example the argument, 'Since our departure, we have driven past a service station every 5 miles, therefore, we will continue to drive by service stations every 5 miles until we reach our destination'. The viability of this argument is determined by the inference made about the actual state of affairs, and thus the fallacy made here can be understood as a generalisation based on insufficient evidence.

It should be noted that even if a fallacy is spotted in an argument, that does not mean the conclusion of the proposed argument must be false, this is a fallacy in itself, called the fallacy-fallacy (Richardson, Smith, & Meaden, 2012). Instead, it may be the case that the inference made or the argument constructed is faulty and a more cogent or valid argument can be constructed to

support the conclusion. This is an important point, since a person may arrive at a correct conclusion incidentally, by employing faulty reasoning. Nonetheless, critical thinking is concerned with the process of thinking that would see the increased likelihood of making an informed decision about the reliability of information, this is what our approach aims to assist people in uncovering.

The work reported here stems from a project with a long term aim to detect fallacies in the media by making use of machine learning (ML) approaches. Based on the output predictions, a tool could thus flag to consumers where fallacies were present and of what kind. However, since fallacies can be expressed in many different ways, the first step in creating this tool would be to create a labelled dataset that could be used to train ML. We now turn to a discussion of early work that has been completed in prototyping this approach.

Method

A design was adopted in which resources (described below) were annotated with markers to the fallacies they contain. It was hypothesised that particular fallacies would occur more frequently across different news topics, which would allow us to develop a prototype tool that would flag these topic-based signature fallacies to consumers. To conduct this preliminary work, two timely issues in Australia were chosen in late 2017 and logical fallacies were systematically recorded and identified in online mainstream media and social media platforms. These two timely issues were: The Safe Schools Program and the Adani Carmichael Mine controversy.

Topic Selection

The Safe Schools Program was an initiative presented by the Safe Schools Coalition Australia that would allow schools to opt-in to utilise educational resources on LGBTQIPA issues that would aim to promote tolerance and understanding in the school setting (Reynolds, 2017). This became a widely controversial topic, as conservative political figures and columnists interpreted the initiative as an "indoctrination program" that would confuse children's understanding of gender and sexuality (Devine, 2017; Kelly, 2016). Moreover, the national plebiscite to legalise gay marriage was nearing and campaigns to counter-act social pressures of voting 'yes' and voting 'no' emerged, making the plebiscite a polarising issue. More significantly, Steve Dickson, the One Nation Queensland leader at the time, had made false claims that the program had contained "highly explicit material... directed at young children." This caused rumours to spread as well as a great deal of concern among conservatives over the wider consequences of legalising gay marriage (Louden & Rowe, 2017).

The second topic came from a highly controversial decision by the Queensland government to support the construction of the largest coal mine in Australia – the Adani Carmichael mine - for the Indian mining mogul, Gautam Adani. This decision received considerable support from some sectors of the Australian community, and extensive criticism from other ones. Those who

supported the construction of the mine, saw the promise of thousands of jobs for locals, and concomitant economic benefits for Australia, as a great benefit. However, these benefits have been met with scepticism, for example the number of jobs proposed by Adani has continued to change and reduce in number (Cox, 2019; Robertson, 2017). Moreover, climate scientists and environmentalists are concerned about the greater environmental impacts of the mine, such as the carbon emissions it will produce, the ongoing utilisation of the Great Artesian Basins and the impact on the Great Barrier Reef (Slezak, 2017). As such, debates around the legitimacy of climate change and the long-term economic benefits of coal mining has proliferated.

Content Sources

The data was collected from multiple online news aggregators, specifically: YouTube, GoogleNews and Reddit. These aggregators were used because they were identified as spaces on the web that organized broad areas of discussion (Hussain et al., 2018; Zannettou et al., 2017). Moreover, both Reddit and YouTube have been understood as spaces where disinformation can be propagated (Hussain, Tokdemir, Agarwal, & Al-Khateeb, 2018; Zannettou et al., 2017). Google News was chosen as it is an increasingly popular online news aggregator, where research has shown a 25% increase in consumers turning to Google for news access, for example there were "1.6 billion visits per week in 2018" (News Media Alliance, 2019, p. 3).

Searching for the chosen topics on YouTube, GoogleNews and Reddit was done systematically. For YouTube, the first 20 results that appeared under 'Adani Carmichael mine' and 'The Safe Schools Program' were analysed. The search was optimised by narrowing the 'Upload date', 'Time' and 'Duration' search filters to 'This month' for recent content, 'Video' was also selected as well as the filter 'Short < 4 minutes', for time management on data collection. Similarly, the first 20 news articles were selected for GoogleNews and Reddit, and the filter 'past month' was selected to produce recent headlines and discussions in the search results.

Due to the presence of multiple hyperlinks in the individual articles, descriptions and comments section, the author performing the coding (GY) followed the links that appeared in the initial article or post for analysis and tagging, but would not follow any subsequent links.

Coding Approach

The first author conducted all the fallacy tagging, using an online annotation tool called Hypothesis. Hypothesis is a web browser extension and open source program that allows the user to highlight text on the web and annotate it for either public or private purposes (a private account was used for this project).

Fallacies were tagged by highlighting an area of text on the page where the fallacy was spotted, and the name of the fallacy was inputted into the annotation box, which was prefixed with a hashtag, allowing for easier categorisation and sorting of the data during the analysis phase. For

example, if the strawman fallacy was spotted in a text or video, a part of the available text was highlighted, and within the annotation box that appeared '#strawman' was entered and saved.

Overall, 600 pieces of data (tagged fallacies) from online news and social media articles were collected and tagged. Figure 1 displays the types of fallacies that were identified in these news sources as well as how often the fallacies appeared. It should be noted, that the logical fallacies that were tagged were not consciously being sought by the author. The data that was collected and eventually aggregated were fallacies that the author had spotted in the text, referencing multiple online tools such as www.fallacyfiles.org and www.yourlogicalfallacyis.com to aid in identifying them.

Results

Over 600 pieces of data (tagged fallacies) from online news and social media articles were collected and graphed (see fig. 1), displaying the kinds of fallacies that appeared in these news stories, and their frequency.

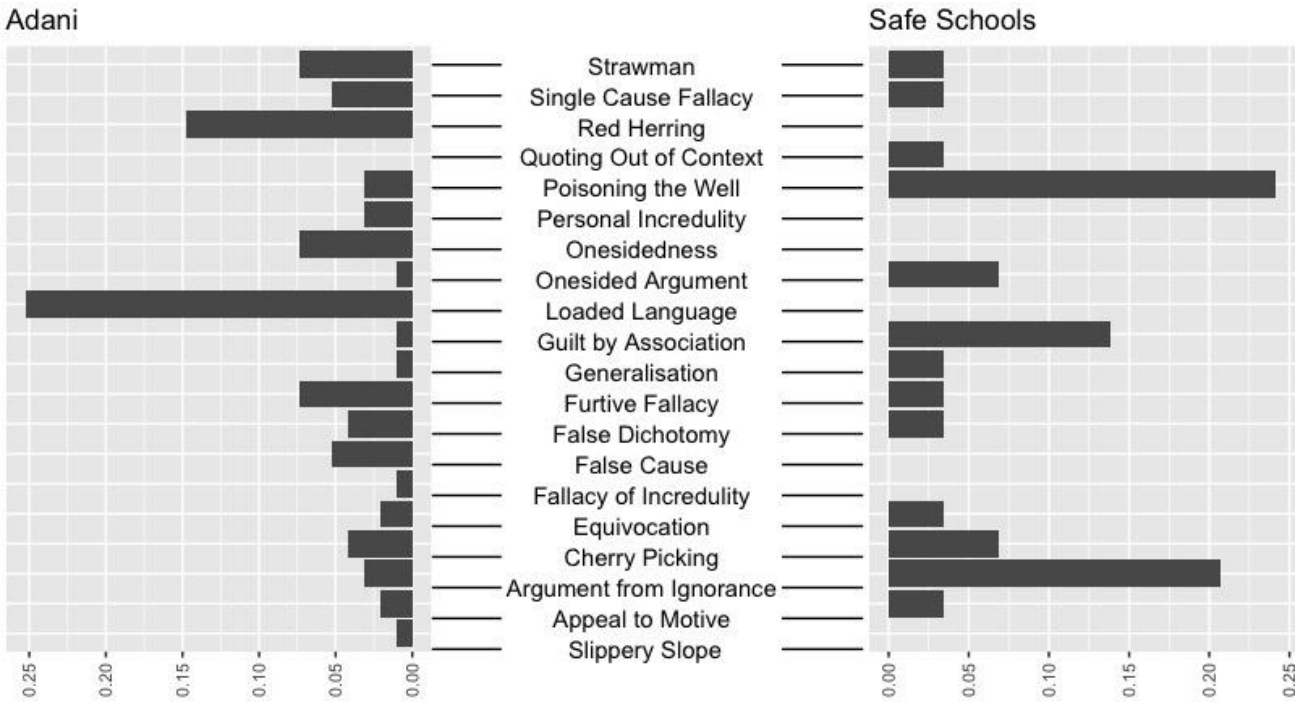

Figure 1: The fallacies tagged using the method described above demonstrating the markedly different signatures for the two topical issues explored in this study. Note that far more fallacies were tagged in the Adani issue than the Safe Schools one, demonstrating that this issue is far more plagued with logical fallacies.

We see from Figure 1 that there are clear differences in the distribution of particular fallacies. In the case of the Safe Schools Program controversy, we see the poisoning the well fallacy as the

most frequently occurring, with the argument from ignorance coming in second. The poisoning the well fallacy is an attempt to discredit opposing arguments, typically by setting up an ad hominem attack and using loaded language. An argument from ignorance, however, is arriving at a conclusion on the basis of the lack of evidence supporting the contrary position. (The fallacy files, 2012). In contrast, loaded language appeared far more frequently when it came to the Adani Carmichael mine issue, with red herring appearing as the second most prominent fallacy emerging from this news story. In the context of argumentation, loaded language is the leveraging of the connotation of words to 'smuggle' in evaluations that have yet to actually be demonstrated. The red herring fallacy, on the other hand, is the rhetorical strategy that aims to mislead one's opponent away from the original issue by introducing an irrelevant topic (The fallacy files, 2012).

This preliminary analysis allowed us to focus and refine the aim of our proposed media literacy tool: to understand and identify the distinct and recurring logical flaws that likely pertain to a particular topic area. We have come to call this strategy 'Fallasigns' (the amalgam of 'fallacy signatures'). We believe that identifying signature fallacies presents a potentially richer and more viable method for teaching people to critically evaluate news in different ways. An automated version of Fallasigns would ideally track the signature fallacies of different news stories and present them to the online user to forewarn them. This approach is further described below.

Fallasigns: Protectionist and Empowerment Literacy

Scholars have identified two major classes that different media literacy strands fall into; the protectionist and empowerment class (Aufderheide, 1993; Hobbs, 1999). The empowerment camp, as described by RobbGrieco & Hobbs (2013) is concerned with giving media consumers the tools to deal with the imbalance of power between media companies and their audiences, this can involve cultivating critical discussion, reflective practice and media production skills. On the other hand, the protectionist camp is concerned with media effects, a perspective that sees media consumers as defenceless against media effects requiring advocacy in the way of media regulation, reform and inoculation.

The strategy that Fallasigns aims to implement can be said to fall into both the protectionist and empowerment camp of media literacy. In one way, to identify fallacy signatures seems conducive to inoculating people from faulty reasoning (and so working to protect them). On the other hand, fallacy signatures provide a means by which media users can be taught to think strategically about the information they find on the internet. More specifically, it provides a tool for evaluating information. This ability is one of the most important skills emerging in the current media landscape, where, as we discussed above, media consumers are confronted with an unparalleled volume of information, which can be produced at ease, and is often strategically intended to mislead news consumers (Albright, 2017). As such, Fallasigns can be understood to reside within an evaluation approach to media literacy. It is an approach that aims to be useful

within multiple pedagogical contexts, as it invites all topics to be analysed for logical errors and can be applied across many forms of media.

While the online version of Fallasigns would operate as a kind of digital heuristic which would relay to the user what fallacies to expect in a news story, it is still incumbent upon the user to make a decision about the reliability of the content, which keeps the user in the loop. This circumvents one of the emerging problems that have been identified in many of the technocentric solutions to the fake news phenomenon, which are often seeking to automate the filtering of information and so tend to work in a black box modality (Pasquale, 2015).

Limitations

As an initial prototyping exercise, the study reported here has a number of limitations. The two most significant limitations lie in the number of subjects that were chosen for study, and how the fallacies were identified.

While the Adani Carmichael mine and the Safe Schools program were timely and topical issues in Australia during the data collection process (The Adani Carmichael Mine still remains topical at this time), only two subjects were chosen in order to provide a small sample of data for this exploratory study. A significant phase of further study is required to work towards a functional Fallasigns product, even at a prototype phase.

As to the process of identifying fallacies, this approach involved one person (the first author) identifying logical fallacies using online resources and their discretion. The potential problem this presents is coder bias, where the coder may be seeing fallacies that aren't present, especially within politically and socially charged content that do not agree with their world view. Multiple coders and checks for interrater reliability would help to avoid this criticism and would be carried out in future work. As a first step towards developing a data set that could be used to train an AI program to locate fallacies in media, this project provides a first step that a media literacy tool could in fact be created. This paper is, thus, a demonstration of a proof of concept; *signature fallacies*, and their potential to inform a new media literacy evaluation tool that can be further developed in future work. We anticipate that such a tool would provide a way to not only make evaluations about the reliability of information, but also develop the ability of media consumers to think strategically about topics that may attract specific logical errors in thinking.

Issues and Future Research

The success of the Fallasigns project, would depend upon several key competencies of the media consumer, for both offline and online strategies. Assuming that the user has basic literacy skills, access to the computer and internet etc., probably the most significant assumption behind this approach is that the consumer will have the ability to see the flaws in arguments that support their view. We know that people are prone to the cognitive biases of selective perception and

selective exposure, which refers to the tendency of people to interpret a message in a way that fits with their underlying worldview, or to consume information that only agrees with that worldview (Severin & Tankard, 2001). These limitations may be avoided if the individual uses the Fallasigns approach to hold their own beliefs to higher standards of reasoning, rather than simply understand it as a way to poke holes in the claims and worldviews of those they do not agree with. However, it will take significant discipline for people to work from this ideal position. Once again, critical thinking and in fact the understanding and application of logical fallacies is about the reasoning process not the conclusion that is reached per se. Another limitation of our strategy, especially in relation to the offline approach, may be the difficulty of calling to mind enough fallacies to select from when trying to determine which are likely to appear in the story, this is our reason for preferring the online approach which can support people in looking up the different fallacies that exist. Though it might be the case that the very exercise the offline strategy offers, may assist in overcoming this potential obstacle.

However, a much stronger objection to this evaluation strategy might be the complexity that the concept of 'logical fallacies' poses to individuals who don't know what fallacies are or how arguments 'work'. Therefore, the ideal iterations of Fallasigns (both online and offline) is one that can incorporate simple explanations and visual cues that work to simplify these concepts. For example, for younger age groups, Fallasigns may be more accessible and effective if gamified or when integrated into a narrative. Lastly, it should be noted that Fallasigns is not intended as a panacea, but as one of many future standardised strategies educators and media consumers would be able to utilise as needed.

Despite the limitations of the signature-fallacy approach explained above, determining fallacy signatures provides a way for media consumers to navigate information in a strategic way, potentially inoculating themselves from being persuaded by seemingly valid arguments, but all the while understanding that the quality and trustworthiness of information can be decided without necessarily having extensive information on any one topic.

Conclusion

While the field of media literacy is fragmented, it is still a field that is growing and has already demonstrated great promise. However, more research is required to demonstrate the effectiveness of emerging media literacy approaches. This paper aims to contribute to the literature in three ways. The first is to introduce scholars and educators to a new way of effectively evaluating information for its likely value and reliability according to signature fallacies. Through the transposition of tools from the philosophy of logic into the media field, we were able to develop a meta-toolbox that holds promise for helping people to achieve a better media literacy. The second contribution is providing ideas for how this new method might be used in both offline and online environments by both educators and consumers. Thirdly, this paper has introduced Fallasigns, the concept of a future platform that would enable further study

into the promise of using fallacies to improve media literacy in our population, laying some of the groundwork for how future research might proceed to develop this approach.

References

Albright, J. (2017). Welcome to the era of fake news. *Media and Communication*, 5(2), 87-89.

Ashley, S., Maksl, A. & Craft, S. (2013). Developing a news media literacy scale. *Journalism and Mass Communication Educator*, 68(1), 7-21.

Aufderheide, P. (1993). *Media Literacy. A Report of the National Leadership Conference on Media Literacy*. Aspen Institute, Communications and Society Program, 1755 Massachusetts Avenue, NW, Suite 501, Washington, DC 20036.

Ávila, J. & Pandya, Z. J. (2013). Traveling, textual, authority, and transformation: an introduction to critical digital literacies. In: Ávila, J. & Pandya, Z. J. Critical digital literacies as social praxis: intersections and challenges, New York, NY: Peter Lang.

Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions: Problems, causes, solutions. *Digital Journalism*, 6(2), 154-175.

Banas, J. A., & Rains, S. A. (2010). A meta-analysis of research on inoculation theory. *Communication Monographs*, 77(3), 281-311.

Bawden, D. (2001). Information and digital literacies: A review of concepts. *Journal of Documentation*, 57(2), 218-259.

Bernardi, C. L. (2016). Commune, the Web, and the Anarchist thought of Mikhail Bakunin. In: Çoban, B., *Social media social movements: The transformation of communication patterns.* London: Lexington Books, pp. 1-19.

Breakstone, J., McGrew, S., Smith, M., Ortega, T., & Wineburg, S. (2018). Why we need a new approach to teaching digital literacy. *Phi Delta Kappan*, 99(6), 27-32.

Brown, J. A. (1998). Media literacy perspectives. *Journal of Communication*, 48(1), 44-57.

Buckingham, D. (2013). *Media education: Literacy, learning and contemporary culture.* Cambridge: Blackwell Publishing Ltd.

Burshtein, S. (2017). The true story on fake news. *Intellectual Property Journal*, 29(3), 397-466.

Chen, D. T., Wu, J., & Wang, Y. M. (2011) Unpacking new media literacy. *Systemics, Cybernetics and Informatics,* 9(2), 84-88.

Cook, J., Lewandowsky, S., & Ecker, U. K. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PloS One*, 12(5), 1-21.

Cope, B., & Kalantzis, M., (2010) New media, new learning. In: Cole, R., D. & Pullen L. D. *Multiliteracies in motion: current theory and practice,* New York, NY: Routledge, pp. 87-104.

Coşkun, İ. O., & Yılmaz, H. (2016). An Introduction to Consumer Metamorphosis in the Digital Age. In Sezgin, E., ed., *e-Consumers in the Era of New Tourism*. London: Springer.

Cox, L. (2019). Adani jobs explained: why there are new questions over Carmichael mine, *The Guardian*. Retrieved from: https://www.theguardian.com/environment/2019/jun/05/adani-jobs-explained-why-there-are-new-questions-over-carmichael-mine

Devine, M. (2017). An epidemic of transgender children is Safe Schools' legacy. *The Daily Telegraph*. Retrieved from: https://www.dailytelegraph.com.au/rendezview/an-epidemic-of-transgender-children-is-safe-schools-legacy/news-story/085d5681f6bc3dae2357302ab2bee227

Fiske, J. (2002). Introduction to communication studies. 2nd ed. London: Routledge.

Fleming, J. (2014). Media literacy, news literacy, or news appreciation? A case study of the news literacy program at Stony Brook University. *Journalism and Mass Communication Educator*, 69(2), 146-165.

Garcia-Ruiz, R., Ramirez-Garcia, A., & Rodriguez-Rosell, M. (2014). Media literacy education for a new prosumer citizenship. *Comunicar*, 22(43), 16-23.

Hansen, M. B. N. M. (2010). New media. In: W. J. T. Mitchell & B. N. M. Hansen, *Critical terms for media studies.* Chicago: The University of Chicago Press, pp. 172-185.

Hobbs, R. (1998). The seven great debates in the media literacy movement, *Journal of Communication*, 48 (1), 16-32.

Hobbs, R., & Frost, R. (2003). Measuring the acquisition of media-literacy skills. *Reading Research Quarterly*, 38(3), 330-355.

Hobbs, R., & Jensen, A. (2009). The past, present, and future of media literacy education. *Journal of Media Literacy Education*, 1(1), 1-11.

Hussain, M. N., Tokdemir, S., Agarwal, N., & Al-Khateeb, S. (2018, August). Analyzing disinformation and crowd manipulation tactics on YouTube. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 1092-1095). IEEE.

IREXdc. (2017, 18 October), Learn to Discern: Citizen Media Literacy in Ukraine and Beyond (4-minute version) [Youtube Video]. Available at https://www.youtube.com/watch?v=K0ogXQBepX0

Islas, O., Arribas, A., & Gutiérrez, F. (2018). The contribution of Alvin Toffler to the theoretical and conceptual imaginary of communication. *Revista Latina de Comunicación Social*, 73, 648-661.

Kellner, D., & Share, J. (2005). Toward critical media literacy: Core concepts, debates, organizations, and policy. *Discourse: Studies in the Cultural Politics of Education*, 26(3), 369-386.

Kellner, D., & Share, J. (2007). Critical media literacy: Crucial policy choices for a twenty-first-century democracy. *Policy Futures in Education*, 5(1), 59-69.

Kelly, J. (2016). Strip funds for Safe Schools 'social engineering': Tony Abbott. *The Australian. Retrieved from: https://www.theaustralian.com.au/nation/education/strip-funds-for-safe-schools-social-engineering-tony-abbott/news-story/08ef31b79e3b3fe0ef0b5cfc6be5389b*

Koc, M., & Barut, E. (2016). Development and validation of New Media Literacy Scale (NMLS) for university students. *Computers in Human Behavior*, 63, 834-843.

Koltay, T. (2011). The media and the literacies: Media literacy, information literacy, digital literacy. *Media, Culture & Society*, 33(2), 211-221.

Lee, S. (2016). Smart divide: Paradigm shift in digital divide in South Korea. *Journal of Librarianship Information Science*, 48(3), 260-268.

Lee, L., Chen, D. T., Li, J. Y., & Lin, T. B. (2015). Understanding new media literacy: The development of a measuring instrument, *Computers & Education,* 85, 84-93.

Livingstone, S., & Helsper, E. (2010). Balancing opportunities and risks in teenagers' use of the internet: The role of online skills and internet self-efficacy. *New Media Society*, 12(2), 309-329.

Louden, B., & Rowe, E. (2017). FactCheck: does the Safe Schools program contain 'highly explicit material'? *The Conversation*. Retrieved from https://theconversation.com/factcheck-does-the-safe-schools-program-contain-highly-explicit-material-87437

Marten, H. (2010). Evaluating media literacy education: Concepts, theories and future direction. *Journal of Media Literacy Education*, 2(1), 1-22.

Marwick, A. E. (2018). Why do people share fake news? A sociotechnical model of media effects. *Georgetown Law Technology Review*, 2(2), 474-512.

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, 60(3), 413-439.

Meyrowitz, J. (1998). Multiple media literacies. *Journal of Communication*, 48(1), 96-108.

Mihailidis, P., & Viotty, S. (2017). Spreadable spectacle in digital culture: Civic expression, fake news, and the role of media literacies in "post-fact" society. *American Behavioral Scientist*, 61(4), 441-454.

Miller, V. (2011). *Understanding digital culture.* London: SAGE Publications.

Mujica, C. (2012). Creating shared dialog through case study exploration: The global media literacy learning module. In: Mihailidis, P. *News literacy: Global perspectives for the newsroom the classroom* New York: Peter Lang, pp. 97-117.

National Literacy Trust. (2018). *Fake news and critical literacy: The final report of the Commission on Fake News and the Teaching of Critical Literacy in Schools*. Retrieved from: https://cdn.literacytrust.org.uk/media/documents/Fake_news_and_critical_literacy_-_final_report.pdf

Nettlefold, J., & Williams, K. (2018). Can you tell fact from fiction in the news? Most students can't. *The Conversation.* Retrieved from: http://theconversation.com/can-you-tell-fact-from-fiction-in-the-news-most-students-cant-102580

News Media Alliance. (2019). *Google Benefit from News Content: Economic Study*. Retrieved from: https://www.newsmediaalliance.org/wp-content/uploads/2019/06/Google-Benefit-from-News-Content.pdf

Park, S., Fisher, C., Fuller, G., & Lee, J. (2018). *Digital news report: Australia 2018*. Retrieved from: https://researchsystem.canberra.edu.au/ws/portalfiles/portal/25520278/apo_nid174861_797956.pdf

Pasquale, F. (2015). The black box society. England: Harvard University Press.

Potter, W. J. (2016). *Media literacy*. 8th ed. London: SAGE Publications.

Reynolds, E. (2017). Myths around Safe Schools and same-sex marriages debunked. *news.com.au*. Retrieved from: https://www.news.com.au/lifestyle/gay-marriage/myths-around-safe-schoolsdebunked/news-story/374b23a137c5c9ac7b63ace15b9806c7

Richardson, J., Smith, A., Meaden, S. & Flip Cretive. (2012). *Thou shalt not commit logical fallacies*. Retrieved from: https://yourlogicalfallacyis.com

Ritzer, G., & Jurgenson, N. (2010). Production, consumption, prosumption: The nature of capitalism in the age of the digital 'prosumer'. *Journal of Consumer Culture*, 10(1), 13-36.

RobbGrieco, M., & Hobbs, R. (2013). *A field guide to media literacy education in the united states*. Retrieved from: https://mediaeducationlab.com/sites/default/files/Field%2520Guide%2520to%2520Media%2520Literacy%2520_0.pdf

Robertson, J. (2017). Adani: Indian miner referred to consumer watchdog over 'misleading' employment claims. *ABC*. Retrieved from: https://www.abc.net.au/news/2017-12-06/adani-jobs-accc-legal-case-chris-mccoomb/9226560

Severin, J. W., & Tankard, J. W. (2001). *Communication theory: origins, methods, and uses in the mass media*. 5th ed. New York, NY: Addison Wesley Longman, Inc.

Shin, J, & Thorson, K. (2017). Partisan selective sharing: The biased diffusion of fact-checking messages on social media. *Journal of Communication*, 67(2), 233-255.

Slezak, M. (2017). Why Adani's planned Carmichael coalmine matters to Australia – and the world. *The Guardian*. Retrieved from: https://www.theguardian.com/business/2017/aug/16/why-adanis-planned-carmichael-coalmine-matters-to-australia-and-the-world

The Fallacy Files (2012). Taxonomy of Logical Fallacies. Retrieved from: https://www.fallacyfiles.org/taxonnew.htm

The News Literacy Project (2020) *Mission*. Retrieved from: https://newslit.org/about/mission/

Van Dijk, J. A. (2005). The Deepening Divide: Inequality and the Information Society. London: SAGE.

Van Dijk, J. A., & Van Deursen, A. J. (2014). *Digital skills: Unlocking the information society*. New York, NY: Palgrave Macmillan.

Vleet, J, E Van. (2011). *Informal logical fallacies: A brief guide.* Maryland: University Press of America.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151.

Wineburg, S., & McGrew, S. (2017). *Lateral reading: Reading less and learning more when evaluating digital information*. Retrieved from: https://pdfs.semanticscholar.org/5b26/9628f4dc29b514dfcb0b5e429e49fc0dae6d.pdf

Woolley, B. (1992). Virtual worlds: a journey in hype and hyperreality. London: Penguin Books.

Zannettou, S., Caulfield, T., De Cristofaro, E., Kourtelris, N., Leontiadis, I., Sirivianos, M., ... & Blackburn, J. (2017, November). The web centipede: understanding how web communities influence each other through the lens of mainstream and alternative news sources. In *Proceedings of the 2017 Internet Measurement Conference* (pp. 405-417).

Dear Professor Mills,

Thank you for your comments, and those of the reviewers, which we have sought to address in the revised manuscript.

Reviewer #1 requested the authors to provide some discussion around new media and how the Fallasign tool can be considered in light of new media literacy. This can be found under ‘Access and Prosumption’.

Reviewer #2 pointed out grammatical errors in this paper, which have since been corrected.

You will see in our response (below) that we have attempted to reply to each comment individually.

Kind regards

The authors (blinded for review)

RESPONSE TABLE

Reviewer #1	
Although much improved, an important consideration the authors have missed in developing their discussion on media literacy is how new media literacies figure in this debate. After all, there is a section entitled 'Working Toward a New Media Literacy Tool'.	The authors have taken this feedback on board, and as such this can be found under the ‘Access and Prosumption’ heading. See page 3.
Adding some further context as to how the tool can be thought in line with 'new media literacy' would nicely complete the discussion.	This is also covered under the ‘Access and Prosumption’ heading. See page 3.
The Access and 'Prosumption' section could be reduced in length for greater succinctness (further notes are available on my reviewer copy)	The authors have taken this feedback on board and as such this section has been reduced.
Reviewer #2	

1."the need for standardised media literacy strategies" - desirable to say need for whom - who is it that the authors particularly have in mind? Who will mainly benefit from the approach described?	This has been amended.
2. Fake news should be phenomenon, not phenomena?	This has been amended.
3. "As information is easier to access" - easier than when or where?	This has been amended.
4. "dealing with evaluation" - specifically, evaluation of what?	This has been amended.
5. "news topics can attract" - desirable to give an example of what kind of news topics - some much more liable to these problems than others	This has been amended.
6. straight forward = straightforward	This has been amended.
7. "high volumes of information is" - should be are	This has been amended.
8. "Emphasis in media literacy training range" - was this in the original quotation, and if so, was it ranges, not range?	This has been amended.
9. "understating" - is this supposed to be understanding?	This has been amended.
10. "(Wineburg" does not need bracket	This has been amended.
11. "decidedly" - is this supposed to be deliberately?	This has been amended.
12. "a great deal of concern over the wider consequences of legalising gay marriage among conservatives" - more	This has been amended.

likely to be "concern among conservatives over the wider"??	
13. "While Google News was" - to make this a complete sentence rather than a sentence fragment, drop "While"	This has been amended.
14. "fallacies that were tagged were not consciously being sought by the author" - agreed, but the paper needs to acknowledge that unconscious bias is possible, and it's desirable therefore for the paper to reflect on the desirability of more than one coder with the aim of working towards inter-coder reliability	This has been amended.
15. "gernasliation" - is this generalisation?	This has been amended. See page 12.
16. The paragraph following Figure 1 has a lot of observations that are not the same as what Fig. 1 reveals. All the discussion in that par about the fallacies that were most evident is different to what is shown in the figure. I don't know how to account for this?	The correction can be found under the figure under page 12 and leading into page 13.
17. "the difficulty of calling to mind enough fallacies" - agreed, but this though is an opportunity to make the point that the exercise provides a useful basis for training readers in how to spot fallacies	This suggestion has been taken on by the authors, and has been actioned on page 15.
18. "might be the complexity of the concept" - should this be "might be the complexity that the the concept"?	This has been amended.