



City Research Online

City, University of London Institutional Repository

Citation: McKay, D., Makri, S. & Buchanan, G. (2024). [citation needed]: An Examination of Types and Purpose of Evidence Provided in Three Online Discussions on Reddit. In: UNSPECIFIED . New York, USA: ACM. ISBN 9798400704345 doi: 10.1145/3627508.3638321

This is the published version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/32540/>

Link to published version: <https://doi.org/10.1145/3627508.3638321>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk



[citation needed]: Characterising Evidence Use in Three Distinct Reddit Communities

Dana Mckay
RMIT University
dana.mckay@gmail.com

Stephann Makri
City, University of London
stephann@city.ac.uk

George Buchanan
University of Melbourne
George.buchanan@unimelb.edu.au

ABSTRACT

In a world where misinformation is abundant, and conspiracy theorists urge others to ‘do their own research’, how do people use evidence in online discussions? What types of evidence do they provide, and for what purpose? Decades of human information interaction research has focused on making it easy to share and discuss information online; and decades of information literacy research have examined how to promote critical thinking and evaluation. However, there is a lack both of systematic analyses of evidence use in online discussions, and the ways community norms affect use of evidence in those discussions. We present a mixed methods analysis of the use of three formats of external evidence (images, links, and direct quotation by using blockquotes) across three Reddit communities with very different norms. One focuses on promoting conspiracy theories, another on debunking them, and a third on personal view change. We investigate the use of these evidence formats within and between communities to understand how evidence is used in different kinds of conversation. Our findings support the design of online information tools that promote good evidentiary practice.

CCS CONCEPTS

• **Information systems:** Information retrieval; • **Human-centered computing:** Human computer interaction (HCI);

KEYWORDS

Reddit, Evidence Provision, Citation, Verification, Factchecking

ACM Reference Format:

Dana Mckay, Stephann Makri, and George Buchanan. 2024. [citation needed]: Characterising Evidence Use in Three Distinct Reddit Communities. In *Proceedings of the 2024 ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '24)*, March 10–14, 2024, Sheffield, United Kingdom. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3627508.3638321>

1 INTRODUCTION

We live in an age where high quality evidence is readily available, yet misinformation remains one of the greatest information challenges of our time. Paradoxically, the call to ‘do your own research’

is more associated with conspiracy theorists than empirical scientists [35]. Online discussions on social media are a common vector for the spread of misinformation [19; 67; 71]. Much research has sought to examine how fast misinformation spreads [62], why people share it [15; 54; 60], and how people on social media persuade each other of true and false claims [42; 68].

One way posters on social media discussions can strengthen their claims is to cite external evidence, which makes their posts demonstrably more persuasive [49]. To provide high quality evidence, however, posters must be able to find and identify it [58; 64]—skills that reduce susceptibility to misinformation [30]. But how are these skills applied in practice? While prior research has studied why online community members share misinformation [15] and how they persuade others [44; 46; 68], what type of content they provide as evidence of their positions and the evidentiary purpose of providing that content is not yet understood. We also do not know whether, or how, technology could better support people to provide high quality evidence.

The way people use online communities has been studied extensively (e.g. [6; 47; 55]). However there is a surprising dearth of research on the use of evidence in these conversations. While argumentation is well understood [42; 44], the use of evidence has only been touched on briefly, and in the context of single communities, [28; 49] rather than providing a more general view, and communities may have very different norms [13].

Of course, the function of evidence in online communities is to persuade, but what does it take to persuade someone? That depends on the claim being made. Research shows that changing our personal views requires evidence that considers our personal situation, but this evidence need not be scientific [9]; technical discussions and decisions may require a more scientific evidence base [26]. Conversely, conspiracy theorists have been accused, possibly unfairly, of ignoring evidence [35], or engaging poorly with low quality evidence [65]. In this situation, evidence may entrench contrarian positions [42], or simply play no role.

Given the differences in evidence needed to persuade, differences in different types of online conversation are likely. But, like the use of evidence in general, this is an open question. We performed a mixed methods analysis of how three common approaches to providing external evidence, via *images*, *links*, and *direct quotes*, are used in three contrasting Reddit communities: r/Conspiracy, r/ChangeMyView and r/DebunkThis. These subreddits host different types of discussion: one focused on separating fact from fiction (r/DebunkThis), one supports people in reflecting on and possibly changing their personal views (r/ChangeMyView), and the third discusses and nurtures conspiracy theories (r/Conspiracy). As these communities share the same technological platform (Reddit), differences can be reasonably ascribed to community norms, rather

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHIIR '24, March 10–14, 2024, Sheffield, United Kingdom

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-0434-5/24/03

<https://doi.org/10.1145/3627508.3638321>

than interface mechanisms. Alongside describing the use of evidence generally, we compare and contrast what kinds of evidence are used and why in each of these communities.

The key contribution of this paper is a typology of information used as evidence, and the way it used (its evidentiary purpose). This typology advances our understanding of the ways evidence are provided in different online communities. These findings have important implications for the design of technology to better support transparent, traceable evidence provision and for setting community norms around evidence provision.

In this paper, we first situate our research in the context of prior work, then describe our data collection and analysis method. Next, we present our findings across the three evidence provision approaches (images, links and direct quotes), comparing and contrasting their use across subreddits. We then discuss our findings in the context of the literature, and offer suggestions for how technology may better support good evidentiary practice, before drawing conclusions and suggesting key avenues for future work.

2 BACKGROUND

In this section, we review the literature on online communities focusing on their social role and the known concerns about how debates unfold within them. Next, we examine prior work on information literacy—the principles and practice of understanding quality information, including evidence. Third, we discuss argumentation and persuasion online. This sets the context for a review of the role evidence plays in online discussions, particularly in combatting misinformation. Finally, we examine prior studies of Reddit to contextualize our method.

2.1 Effective and Healthy Online Communities

Online communities have been the subject of information interaction research for over 20 years [47]. Early research defined online communities as any online space where people meet to share information or support, or have social interactions. Communities are made up of the underlying platform, the people who post, the posted content, moderators [18; 47], and—arguably—lurkers [43], who read but do not post. It is vital to consider lurkers who may be influenced by the ideas in online communities; research on misinformation has shown that consensus in online communities may incline readers to believe ideas that are not true [16].

Effective online communities have good *sociability*—they are places where positive social interactions take place. In her key work on online communities, Preece identified three elements of good sociability [48]: having a *community purpose*, having *good people* (including both moderators and participants), and having *strong and clear policies*. These policies support community norms—that is, what is deemed acceptable practice for a community. Norms may be implicit or explicit, and address behaviour (such as frequency of posting) and the content of posts [13; 33].

Preece also noted the problem of usability [48]—having tools that people can easily use to meet their own aims and do what is necessary to abide by community norms. Considerable research has gone into understanding how to support online communities by designing usable interfaces, and by reducing the burden on administrators and moderators in upholding norms. Key design

features include limiting anonymity [72], community moderation (including flagging, upvotes and downvotes [40; 63]), labelling bad behaviour [61], muting other posters [25], and sharing the rules [38] to support new users. Even with good sociability and clear community norms, though, online communities may promote positive behaviour (such as engaging with health professionals) or negative behaviour (such as bingeing and purging) [11].

The dark side of online communities has long been acknowledged [43]. They can promote unhealthy [11] or antisocial behaviour [56], and many fear they are a conduit for the spread of misinformation [17; 27; 60; 70]. The proliferation of special interest groups has made online discussion groups places where people can seek out and prioritise views aligned to their own [22; 57; 66] and created numerous fora where people try to persuade others of their views [40]. This persuasion can sometimes be effective, especially in communities where a single message is allowed to dominate [46].

Communities can protect themselves from misinformation. Communities with polite but persistent disagreement can prevent people from believing misinformation [10; 16; 36], as can those that normalise fact checking and robust evidence provision [28]. What is the role of community norms in supporting (or discouraging) evidence provision? How do these norms affect the use of tools to support the provision of evidence? Reddit, discussed further in section 2.5, with its consistent interface but wildly variant communities is an excellent place to study these questions.

2.2 Information Literacy

In a world of widely available information and rampant misinformation, information literacy has been identified as a cornerstone in the fight for accuracy [1; 17; 53]. In contrast with other literacies (e.g. media literacy) information literacy has been empirically demonstrated to improve recognition of misinformation, [30; 31].

The demonstrated effectiveness of information literacy principles is possibly because evaluating information is a key skill [1]. This skill applies not just to assessing information, but also to understanding where it comes from, both in terms of its creator, and the information-seeking process that led to finding it [64]. Reflecting on information creation aligns with ‘inoculation’; strategies for combatting misinformation [21].

One simple, widely-used information evaluation technique is the CRAAP (acronym intentional) test [2], which supports evaluation of the *c*urrency, *r*elevance, *a*uthority, *a*ccuracy and *p*urpose of information. Most relevant to our study are *authority* (is information from an authoritative source?), *accuracy*, which is notoriously difficult to assess [1], and *purpose* (why the information has been created, which may reveal inherent biases). Good evidence is *accurate*, *authoritative*, and fit for *purpose*.

While we know information literacy skills reduce susceptibility to misinformation, there has been little study of whether, or how often, people provide the evidence information literacy guidelines recommend they should in everyday online discussions. Online discussions are an important context, as they can catalyse view change [39]. In a world where misinformation is rife, it is vital that view changes are underpinned by high quality evidence. As good evidence on its own is often unpersuasive [35; 36], we next discuss how people argue and persuade, before returning to evidence itself.

2.3 Argumentation and Persuasion

Evidence is of little use unless it is persuasive, but traditional models of persuasive information (consistent and overwhelming messages [46]) do not apply on the internet, where the scope for single messages to dominate is limited [8]. The internet has made it easy for people to ‘do their own research’ (however ineffectively) [35], allowed people to seek and prioritise views aligned to their own [22; 57; 66], and created numerous fora where people try to persuade others of their views [40].

In this context understanding persuasion and argumentation is both more possible [32; 44], and more important than ever. Exposure to diverse views (such as through argumentation) may be a strong defence against polarization and misinformation [29] (though it may not, too [42]). There is however, compelling evidence that at least some people engage with views other than their own [5; 40; 69], sometimes this resulting in view change [39].

Considerable work has been dedicated to identifying and classifying arguments and viewpoints online (e.g. [20; 52]), and understanding the persuasiveness of different types of argument [10; 68], and the role of psychology in avoiding backfire effects [36]. While we know that argumentation and persuasion play an important role in online discussions, especially in influencing peoples’ views, we do not yet have a clear picture of how evidence is used to argue and persuade in online discussions.

2.4 Evidence’s Role in Online Communities

To understand the role of evidence in online discussions, we must first define it. Evidence is defined in the Oxford English Dictionary as ‘grounds for belief’, specifically ‘facts or observations adduced in support of a conclusion or statement’ [45]. What does it mean for something to be grounds for belief, though? Goodnight [26] divides contention into three spheres based on the type of evidence needed to reach a conclusion:

The **personal** sphere, needing evidence only to satisfy oneself;

The **technical** sphere, wherein evidence must meet standards agreed on by a practice, or profession;

The **public** sphere, where evidence must be sufficiently comprehensive to satisfy the public good.

We have already seen that some online communities engage in good evidence practice [28] and that evidence is persuasive [49]. If this were the whole picture, we could conclude that merely using evidence would avoid the worst effects of misinformation, and thus normalizing evidence use would address the knock-on problems created by believing in misinformation. However, evidence alone is not enough: some people who hold alternative beliefs appeal to scientific norms (e.g. randomized controlled trials) while rejecting the material evidence [9], others simply reject evidence after evaluating it extensively [39]. This differential response to evidence makes the role of community norms around evidence an interesting topic of enquiry. Taking Wikipedia as a place where a community gathers to create knowledge some research has shown that citations were added in response to controversy [23]; links to authoritative evidence were preferred, and links to low-quality evidence were often deleted. Moderator interest can also promote the inclusion of evidence [14]. Arguably, though, Wikipedia is a single community

with a single standard for evidence. What happens when communities have the same tools, but different norms around evidence? This research seeks to answer this question.

Given its importance, there is surprisingly little research on the role of evidence in online discussions. While the use of argumentation and evidence have been investigated in the education domain—to understand what contributes to strong pedagogical outcomes (e.g. [37]), there is little research beyond this domain. One study focusing on a Reddit political discussion found that <7% of posts cited external evidence through hyperlinks, 9% used political experience, and 20% used hypothetical situations in argumentation. This study did not examine other evidence formats (e.g. quotes). Another study investigated evidence provision in one of the subreddits we examine in this study; r/ChangeMyView—an online community where posters ask others to provide arguments (often including evidence) to change their personal viewpoints [44]. This study found extensive evidence was provided when attempting to change views on social or moral issues, but that this had little impact on beliefs., but that more evidence was more likely to affect beliefs. These studies offer glimpses into a complex field: while evidence may persuade, we do not yet know what kinds of evidence people provide, nor what their aims are in providing it.

2.5 Research on Reddit

Reddit is often used for online community research [44; 53] due to being a large publicly-available dataset encompassing many communities with widely varying norms [15]. Reddit also has a storied reputation, as ‘cesspool’ [26], ‘part of the propaganda pipeline’ [2] and a place that ‘makes people think’ [61].

Community norms are particularly important on Reddit, existing at both pan-Reddit and subreddit levels [13]. Moderation on Reddit is effective in improving behaviour, even among previously hateful posters [59], and in maintaining Reddit-wide and subreddit-specific community norms [12; 13; 24]. Reddit users also seem to alter their behaviour when posting on different subreddits [51]

While most of the research on Reddit is pan-Reddit [50], there are prior studies of all three subreddits that we examine. R/ChangeMyView is one of the most studied subreddits; mostly quantitative analyses have examined the mechanics of persuasion, revealing that certain types of language are more persuasive [68], as is the inclusion of evidence in a post, and that most original posts (OPs) contain both claims, but not evidence [44]. R/Conspiracy analyses have demonstrated clear linguistic differences from mainstream subreddits [32; 55] and that ‘dramatic’ events, such as terrorist attacks or mass shootings, bring in new long-term community members [52]. We found only one study of r/DebunkThis; a qualitative analysis describing several argumentation strategies used in debunking spurious claims, including appeal to external evidence [28].

While Reddit is a popular study site [41], there is limited mixed methods research [50], and very little that compares community norms between subreddits (we found only a single, largely quantitative study [11], not focused on evidence). No research, to our knowledge, examines community norms around evidence provision, nor patterns of evidence provision generally. These questions are the focus of our study.

3 METHOD

In this section we explain and justify our data collection and analysis approaches and discuss ethical considerations.

3.1 Data Collection

We used Reddit as it is home to many distinct communities hosted by one technology, allowing us to understand differences in community norms without interface effects. We selected three contrasting subreddits as their community rules and guidelines promote different kinds of evidence. Our data collection was done in 2022, prior to current restrictions to the Reddit API and terms of use.

R/Conspiracy describes itself as a ‘thinking ground’ where all views are respected. It aims to spark discussion on ‘issues which have captured the public’s imagination, from JFK and UFOs to 9/11.’ This is a site of ‘outlaw’ discussion [4] that does not adhere to evidentiary norms such as encouraging verifiable information, focusing instead on whether topics are conspiracy-worthy.

R/DebunkThis markets itself as ‘an evidence-based subreddit dedicated to taking an objective look at questionable theories, dodgy news sources, bold-faced claims, and suspicious studies.’ Posters must explicitly state the claim they would like debunked, providing at least once source to the claim. This is a community dedicated to seeking technical evidence [27] to ‘debunk spurious hogwash’.

R/ChangeMyView is ‘a place to post an opinion you accept may be flawed. . . to understand other perspectives on the issue.’ Original posts are deleted if the poster demonstrates unwillingness to change their view, and replies must either ask the original poster for clarification, or attempt to persuade the poster of an alternative view. This is a community dedicated to affecting the personal sphere [27] and permits (but does not explicitly require in its rules and guidelines) the provision of evidence to support posters’ attempts to change the original poster’s view.

These subreddits therefore all allow posters to cite evidence to persuade others of their claims and arguments, but in different ways; *r/Conspiracy* involves persuading others that a topic might be conspiracy-worthy and encourages posters to ‘keep an open mind’, *r/DebunkThis* involves persuading others that a particular claim might be untrue and actively encourages use of evidence to debunk myths, and *r/ChangeMyView* neither encourages or discourages citing evidence when convincing a poster to reconsider their views, but emphasises healthy debate.

We used a time-bound dataset to ensure comparability across subreddits; we manually captured 50 threads on each subreddit at a fixed timepoint. We captured the most active 50 threads as this is the default view on Reddit, thus capturing the evidence most likely to be seen by visitors to the site. We capped our analysis at 50 threads per subreddit to provide breadth and depth within a manageable qualitative dataset. We downloaded the content of each thread at the moment of capture, to ensure that the versions used had a consistent time-point. All original links and data were retained. For each thread we manually recorded the title, the URL of the post, the number of links to external evidence in the original post, and the number of replies at the time of collection.

Based on an initial exploration of posts across our subreddits, we identified three common approaches to including material from outside the post in both Original Posts (OPs) and replies. The first

was: *Linking* to external sites; second was embedding or linking to *images* that serve as evidence and finally directly quoting text from the OP, previous posters, or other sources by using ‘*blockquotes*’ (indented quotations). Blockquotes are the in-built function Reddit provides for quoting text from another post.

Our data collection approach focused on examining the use of these approaches to evidence provision across the three subreddits, within OPs and responses. All three approaches involve more effort than simply typing text; users had to interact with Reddit’s interface features to embed an image, link or blockquote. Furthermore, as these approaches are automatically detectable, we were able to build a Python-based scraper to identify posts (both OPs and replies) that contained at least one image, link or blockquote. This enabled systematic capture at scale and allowed us to re-verify the manually gathered data that we initially collected.

3.2 Data Analysis

We analysed the dataset via a mixed methods approach. First, we did a qualitative analysis of the use of each evidence format (images, links and blockquotes). The dataset of posts containing links and blockquotes was large, with over 1000 instances each across the three subreddits. To create a manageable dataset, we then randomly selected threads from each subreddit, examining OPs and replies for links or blockquotes. We continued this approach until we had 60 posts per subreddit containing quotes and 60 containing blockquotes. As an example for blockquotes this involved examining 19 threads in *r/Conspiracy* but only 8 in */r/ChangeMyView*. Only 37 images were used as evidence across all subreddits, so these were all analysed. To be included in our dataset, an external item had to be used as *evidence* to support an implied or explicit claim. Items for which we could not identify a claim were discarded.

Once we had formed datasets, we used the general inductive method for analysis. This involves repeated re-reading of the posts in the dataset to identify consistent patterns. The qualitative codes created during this process were named, merged and split iteratively until the analysis reached stability. Each individual image, link or blockquote was treated as a separate datapoint, and was examined within the context of the original post that contained it. We analysed each datapoint to identify the type and purpose of each datapoint. For images, *type* captured the form of image (e.g., photo, diagram, screengrab). For links, *type* captured the nature of information linked to (e.g., news media article, video). For blockquotes, *type* varied according to the source: another poster, or an external source. *Purpose* captured how the datapoint was used to provide evidence (e.g., to appeal to authority, provide facts to support a claim). The authors discussed all types and purposes in detail to reach agreement, assess edge-cases and identify consistent patterns between datapoint types and subreddits. We found that there were more differences than similarities between types between evidence formats. All coding was done systematically and exhaustively to ensure a concrete foundation for the quantitative analysis.

Once all datapoints were qualitatively coded, we *quantitatively* analysed the prevalence of each format (*images*, *links*, *blockquotes*) using chi-squared tests. While the types and purposes found were not directly comparable between these formats, we quantitatively analysed the use of each format between subreddits.

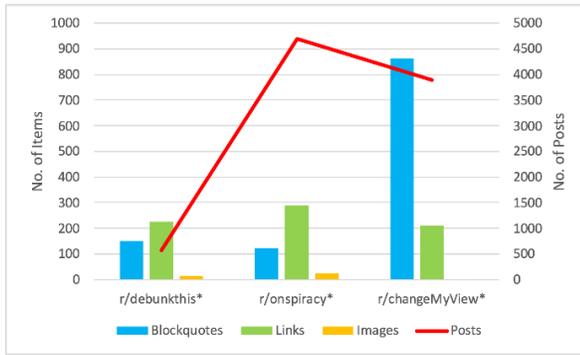


Figure 1: Distribution of images, links and blockquotes by subreddit; total posts per subreddit on the right axis. * = statistically significant ($p < 0.05$)

3.3 Ethical Considerations

This type of analysis is actively excluded from ethics committee oversight at all three authors' institutions. We report our data in line with best practice [7; 50]. Although most Reddit users do not use their real names as usernames and it is likely to be difficult to identify them from their posts, we took steps to further ensure their anonymity; we do not report posters' usernames, nor the specific date and time of data collection. We also paraphrased text quoted from posts to avoid direct identification.

4 FINDINGS

In this section, we first provide an overview and quantitative comparison of the use of evidence formats across r/Conspiracy, r/DebunkThis and r/ChangeMyView. We then report qualitatively on the use of these approaches of evidence provision in each of the three subreddits individually. As the topics discussed in these online communities, including the threads we examined, could be distressing, we advise reader discretion.

4.1 Comparing Subreddits

In this section we quantitatively compare the use of evidence between subreddits. We discuss our qualitative findings in the subsequent finding sections. We performed a chi-squared test to assess the relative frequency of image, link and blockquote use between the three subreddits. This was strongly significant ($p < 0.0001$, $df = 4$, $\chi^2 = 1557.91$). These evidence **formats** were markedly differently distributed across the three subreddits, as shown in Figure 1. For example, blockquotes were mostly found on r/ChangeMyView.

The number of images used for evidence differed significantly between the subreddits ($p < 0.001$, $df = 2$, $\chi^2 = 14.89$). The distribution of links between subreddits was significantly different ($p < 0.001$, $df = 2$, $\chi^2 = 763.53$), with r/DebunkThis having markedly more links than the others. Finally, blockquotes were unevenly distributed across the different subreddits ($p < 0.001$, $df = 2$, $\chi^2 = 745.52$).

In terms of **types** within each of the three formats, the distribution of image types (e.g. photograph, diagram) was too sparse for analysis; images were rare in r/ChangeMyView. The types of linked content (e.g. news, social media) were markedly different between

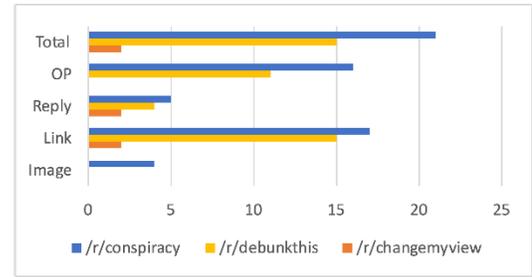


Figure 2: Use of images across subreddits.

subreddits ($p < 0.001$, $df = 24$, $\chi^2 = 91.13$). Blockquotes types (e.g. quoting another poster, quoting a news article) also varied markedly between subreddits ($p < 0.001$, $df = 10$, $\chi^2 = 49.64$) even though quotes of another poster were dominant in the dataset generally.

Finally, we compared the **purposes** of evidence provision across the three formats. Images showed significant differences ($p = 0.0062$, $df = 6$, $\chi^2 = 18.02$) between subreddits. Link purposes were also unevenly distributed ($p < 0.0001$, $df = 18$, $\chi^2 = 75.47$), as were blockquotes ($p < 0.001$, $df = 12$, $\chi^2 = 355.38$).

In short, the subreddits differed markedly from each other.

4.2 Images

There were only 37 images used as evidence across 22 threads in our data set, so we analysed them to exhaustion. We discarded one image not used as evidence, one that had been deleted, and seven images that were being used exclusively as links (the images did not provide any information).

Images were used very differently to provide evidence across the three subreddits (see Figure 2). They were used sparingly and illustratively in r/DebunkThis and r/ChangeMyView. In r/ChangeMyView, they were only used in replies—to help illustrate a different viewpoint. Similarly in r/DebunkThis, images were typically used to illustrate rather than prove. Interestingly, the r/DebunkThis data was skewed by a single thread containing 12 images: nine in the OP and three in a single reply—and the topic was widely regarded as a conspiracy theory: that the US government were behind JFK's assassination, making it more like a thread in r/Conspiracy.

4.2.1 Image Types. A variety of types of image were used as evidence (Figure 3). By far the most common was **screengrabs**. Some sources were identifiable, e.g. a BBC News tweet announcing Boris Johnson's impending resignation. Others were not, e.g. in a post about JFK's assassination a link was given to a screengrab, but neither the image nor the post linked the source.

The next most common type was **photographs**, depicting everything from destruction of the Georgia Guidestones to a man who transformed himself into a lizard. Closely related to photographs was **video stills**, captured from news or other video. A single r/DebunkThis thread about the JFK conspiracy accounted for all **diagrams** and **medical images**, used to debate the trajectory of an assassin's bullet through JFK's body. **Graphs** were only used in a single r/Conspiracy thread about climate change, and **montages** (comprised of multiple images) appeared on a r/Conspiracy thread about Hummer-owner Arnold Schwarzenegger posing with

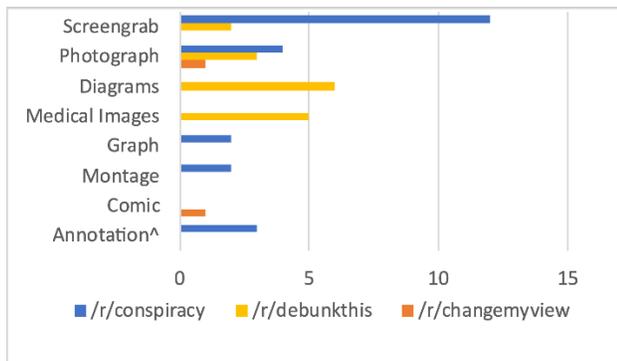


Figure 3: Types of images used as evidence. ^Annotations always co-occur with other image types.

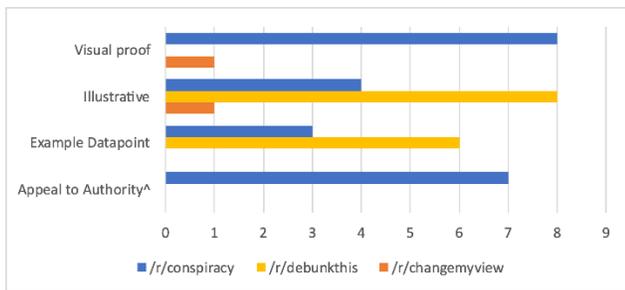


Figure 4: Evidentiary Purposes of Images

environmental activist Greta Thunberg to greenwash his reputation. Finally, a **comic** was shared in r/ChangeMyView to illustrate definitional differences between equality and equity.

Annotation cut across the other image types. Three images in r/Conspiracy were annotated: Two (a photo and screengrab) had roughly-drawn circles added to highlight part of the image. The third annotated a graph with arrows and writing to suggest that climate change is cyclic.

4.2.2 Evidentiary Purposes of Images. Images were used in three different ways to provide evidence (see Figure 4). The simplest was as **visual proof** of a claim (e.g. screengrabs of a BBC News Tweet to prove British Prime Minister Boris Johnson intended to resign). In some screengrabs, images were used as both claim and proof (see Figure 5), which comprised the entire OP in a thread entitled “what the actual f***?” where an image depicting an empty field that had been the site of the Georgia Guidestones both claims and demonstrates they have been demolished.

Another use of images was as an **example datapoint**, providing a single (visual) example to support a more general claim, e.g. medical images to claim JFK could not have been shot at the angle given in his autopsy.

A final use of images was as **illustrative** of a claim, e.g. using a comic illustrating the difference between equity and equality to argue that treating everyone equally is unfair, or graphs showing how the earth’s climate has changed over time to support the argument that climate change is cyclical.



Figure 5: Claim and proof of the Guidestones’ demolition

Visual appeal to authority cut across other evidentiary purposes. This appeal could be explicit, such as the image of a BBC News Tweet, or implied, as with a graph that was presented in a scientific format but lacked any information to assess its veracity. At least one appeal was demonstrably false: a video still containing a bogus Georgia Bureau of Investigation logo. Most appeals were to organisational authority: mainstream news sites or aggregators; two cases appealed to sites that were explicitly Alt-Right. Two images used the authority of public figures, e.g. Robert F. Kennedy.

4.2.3 Summary and Analysis. Images were used very differently to provide evidence across the three subreddits. They were used sparingly and illustratively in r/DebateThis and r/ChangeMyView. In r/ChangeMyView, where images appeared only in replies to help illustrate a different viewpoint. Similarly r/DebateThis, images were typically used to illustrate rather than prove. Interestingly, the r/DebateThis data was significantly skewed by a single thread containing 12 images: nine in the OP and three in a single reply—and the topic was widely regarded as a conspiracy theory: that the US government were behind JFK’s assassination, making it more like a thread in r/Conspiracy.

In contrast with the few images in r/DebateThis and r/ChangeMyView, nearly a third of OPs in r/Conspiracy contained an image, usually as the post’s key message. Appeal to authority was also more common in r/Conspiracy, perhaps reflecting posters’ awareness that their views are alternative, and so need support. This appeal may also be an attempt to use the tools of authority against dominant narratives, and reflect the general approach of r/Conspiracy posts to encourage people to ‘see for themselves’.

4.3 Links

Links were used in a wide range of ways across subreddits to provide evidence. Most links to external sites were found on r/DebateThis, where just under a third (33.1%) contained them. In contrast, links were found in only 6.5% of posts on r/Conspiracy, and 4.7% of posts on r/ChangeMyView.

4.3.1 Link Types. Posts linked to a wide range of content, from Wiki articles to academic publications (see Figure 6; the maximum possible number of links was 60 in each subreddit).

News media articles were the most common content linked to. R/DebateThis and r/ChangeMyView typically featured links to

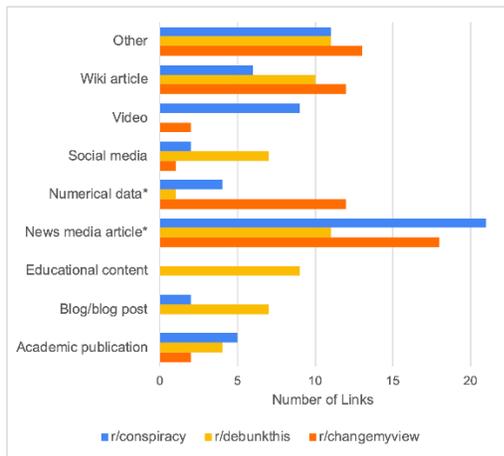


Figure 6: Types of linked content used as evidence. * = statistically significant difference ($p < 0.05$)

trustworthy news media articles. In r/Conspiracy, article credibility varied considerably, with some containing demonstrably false information. Others were factually correct but still could mislead: e.g., an article published by AlterNet (a 'mixed reliability' source: Ad Fontes Media Bias): 'What the Cellphone Industry Doesn't Want You to Know About Radiation Concerns' was cited to support a claim that cell phones cause cancer. Some posters in r/Conspiracy used reliable articles to fuel speculation (e.g., that MKUltra, a 1960-70s CIA human experimentation program, still exists).

In r/DebateThis and r/ChangeMyView, news was used more conventionally, e.g. to explain how a calculation error caused a French presidential candidate to apparently 'lose' votes already counted on a TV graphic. In r/ChangeMyView, news was cited to persuade. For example, a poster cited a WTTW article showing the 1994-2004 Federal ban on assault weapons reduced deaths, to convince the OP who was opposed to gun restrictions.

News articles in r/Conspiracy were mostly cited to *provide facts* supporting claim (12 of 21 links). Source credibility, especially in r/Conspiracy, varied considerably and the reliability of 'facts' must be queried. R/ChangeMyView posts often cited news to *provide numerical data*: 8 of 18 news links pointed to data, e.g. donation data from a Washington Post article suggested some corporations embrace social movements for their own benefit. Numerical data was less common in the other subreddits—c. 9%.

Wiki articles were the next most common type of link. Linking to Wiki articles was more common in r/DebateThis and r/ChangeMyView than r/Conspiracy (17%, 20% and 10% of links respectively). Wiki articles were cited to provide numerical data or facts to support a claim, e.g. an r/Conspiracy post linked to a Wikipedia 'list of largest pharmaceutical settlements' citing numerical data that Pfizer paid a \$2.3 billion settlement for violating the False Claims Act, consequently reasoning Pfizer cannot be trusted. Wiki articles were also used to provide context e.g. a poster in r/DebateThis cited Wikipedia's article on Brandolini's Law that, according to the article, "*emphasises the effort of debunking misinformation, in comparison to the... ease of creating it.*"

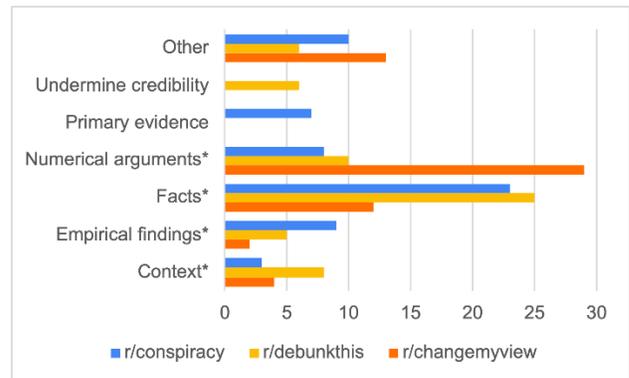


Figure 7: Evidentiary purpose of links. * = statistically significant difference ($p < 0.05$)

Academic publications were cited in all three subreddits, usually to provide empirical findings or numerical data to support a claim. The key difference between subreddits was in r/ChangeMyView credible scholarly work was used to support arguments for view change. In r/DebateThis less credible work was questioned (to support debunking), e.g. one thread questioned the sample sizes in academic articles by a psychologist whose studies appear to validate astrology. In r/Conspiracy even *credible* research was questioned (when it undermined a conspiracy theory), e.g. a poster cited a finding that unvaccinated people who caught COVID were more likely to develop myocarditis than vaccinated people. To cast doubt on this study, another poster noted the authors were members of vaccine advisory groups, with a vested interest.

Numerical data was seen mostly in r/ChangeMyView, but accounted for nearly half of the links in this subreddit.

4.3.2 Evidentiary Purposes of Links. We identified several ways links were used as evidence (see Figure 7).

Links were used in all subreddits to **provide facts**, though least frequently on r/ChangeMyView. In r/DebateThis and r/ChangeMyView, facts were generally accepted in discussion, but in r/Conspiracy some 'facts' were contentious. Even seemingly innocuous posts, such as a request for names of people with similar names to crimes they committed, fuelled speculation: a poster linked to a Wiki article, which said the 'Meaningful Name' TV trope "*has a direct, barely-hidden meaning... that tells you an important thing about the character.*" The poster then claimed the CIA often inserts puns into cover identities they create for people in the media spotlight.

Tables, graphs, and other numerical data used to make **numerical arguments** that comprised almost half of r/ChangeMyView links. In r/DebateThis and r/ChangeMyView, posters usually provided accurate data from authoritative sources while in r/Conspiracy, data was less authoritative: e.g. a poster cited OpenVAERS—an open data source containing CDC data of post-vaccine adverse events—to claim vaccine side-effects are more common and dangerous than stated by another poster. While OpenVAERS data may appear credible, the CDC website states "*a report to VAERS does not mean that a vaccine caused an adverse event.*"¹

¹<https://www.cdc.gov/vaccinesafety/ensuringsafety/monitoring/vaers/index.html>

Empirical findings, mostly from scientific studies, were cited in all subreddits. In r/DebateThis and r/ChangeMyView, posters used empirical findings to persuade. For example, an article in the journal *Renewable Energy* was cited to refute a claim that windmills are environmentally damaging. In r/Conspiracy, empirical findings were often reported second-hand (e.g., by linking to news articles rather than academic papers). This sometimes distorted findings: a news article from LifeSite (banned from some social media platforms for spreading COVID-19 misinformation) selectively reported findings from an academic study on myocarditis and COVID. Rather than report that myocarditis risk is extremely low among both vaccinated and unvaccinated groups, the LifeSite article concluded “*the potential for vaccine-related myocarditis among young males undermines the public health establishment’s persistent refrain that the benefits of vaccination far outweigh any harm.*”

R/Conspiracy contained all seven instances of **primary evidence**, including video footage of the Georgia Guidestones bombing. The timestamp of the explosion (04:33:33am) was used to support a theory that the stones were damaged because they were ‘satanic’; the Centre for Biblical Studies at Midwestern Seminary claims it is likely Jesus was crucified on April 3, AD33².

Using links to **provide context** to an argument or source was most prominent in r/DebateThis. One r/Conspiracy poster used authoritative information from the BMJ and WHO to provide context to their claim it was known that the AstraZeneca and Johnson & Johnson COVID vaccines can cause myocarditis.

All six cases of **undermining the credibility** of (another poster’s) evidence were in r/DebateThis: e.g. in response to an OP with a news article from CBN (Christian Broadcasting Network) claiming a US Family Planning Bill legalises infanticide, a response linked a Media Bias report that CBN has ‘low credibility’.

Links were also used for **other evidentiary purposes**. In a very few cases, links also *appealed to authority*, pointed towards *information sources*, provided *definitions*, *example datapoints*, *examples*, *interpretation*, or *pre-emptive evidence* (in anticipation of another poster rebutting a poster’s claim) and *suggested inaccuracies* in other posters’ evidence.

4.3.3 Summary and Analysis. *News media articles*, *Wiki articles* and *numerical data* were the most common types of linked evidence and featured in all subreddits. However, they were used differently, depending on the subreddit’s remit, serving to fuel conspiracy theories in r/Conspiracy, debunk myths in r/DebateThis and change posters personal views in r/ChangeMyView. Some link types were predominantly found in certain subreddits; social media and blog-posts were most common in r/DebateThis, while video mostly appeared in r/Conspiracy. This suggests posters select the type of linked source to match local subreddit expectations. For example, the preponderance of numerical evidence on r/ChangeMyView and the emphasis on facts from news media on r/Conspiracy may reflect different roles of evidence in these respective communities.

Links to evidence were most often used to provide numerical data and facts, and to provide empirical findings and primary evidence. While in r/ChangeMyView and r/DebateThis these purposes promoted robust discussion underpinned by credible information, in

²<https://cbs.mbts.edu/2020/04/08/april-3-ad-33-why-we-believe-we-can-know-the-exact-date-jesus-died/>

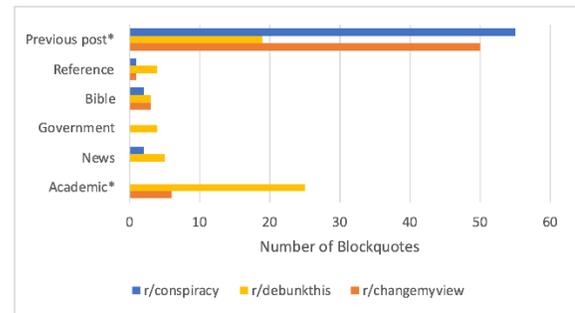


Figure 8: Types of sources cited as evidence using blockquotes. * = statistically significant difference ($p < 0.05$)

r/Conspiracy, they were primarily used to portray alternative beliefs as plausible.

4.4 Blockquotes

Evidentiary blockquotes were found in only 3% of r/Conspiracy posts, versus 22% in r/ChangeMyView and 26% in r/DebateThis.

4.4.1 Blockquote types. There was a clear division between blockquotes that quoted **previous posts**, which link to the original content automatically, and those that provided a clear attribution to external sources. Posters referred to external sources in different ways when quoting from them using blockquotes. Sometimes they simply attributed the source (e.g., ‘Lenin’ when quoting from an essay of his). More often, they provided links or scholarly references. The most common type of external source (Figure 8) was **academic publications**—usually published papers, but sometimes **textbook content** and **literary biographies**. R/Conspiracy quoted no academic sources. Other sources included **news articles**, from mainstream media, Reuters or other independent online sources; extracts from the **Bible**, **reference sources** such as encyclopedia and **government legislation**.

4.4.2 Evidentiary Purposes of Blockquotes. Posters typically blockquoted previous posts in the same thread to re-use putative facts to support an argument. When quoting extracts from external sources, purposes varied. As with links, we noted purposes such as *providing facts*, and *providing numerical data*. There was frequent *provision of primary evidence* and *appeals to authority*, plus some limited reference to *definitions*.

Primary evidence was taken from academic publications and government legislation. Quotes were often accompanied by an explanation of its relevance: e.g., one r/DebateThis poster explained the importance of source credibility and reliability, noting the OP’s source was well-known for pseudo-scientific studies, and quoted text from the American Psychological Association and the Southern Poverty Law Centre to debunk the article.

Blockquotes were also used to **contextualise**. For example, a poster responded to criticism of the use of wind turbines by quoting it for context, then blockquoted two sources of authoritative evidence to debunk it.

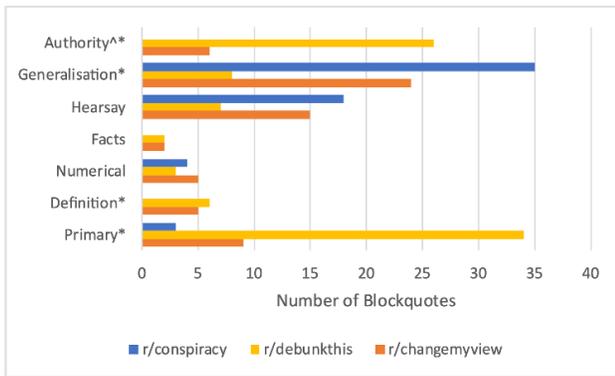


Figure 9: Evidentiary purposes of blockquotes. Key: Authority = to appeal to authority, Numerical = to provide supporting numerical data, Primary = to provide primary evidence etc. ^ = Authority co-occurred with all other purposes. * = statistically significant difference ($p < 0.05$)

Blockquotes sometimes contained data used to make **numerical arguments**, e.g. from academic study that found astrologers' predictions performed no better than chance.

A common use of blockquoting was the provision of '**facts**', all but two without an authoritative source. In Figure 9, we distinguish between **sourced facts**, **generalisations**, and **hearsay** (reposting other posters' claims). Many of the latter two were inaccurate. For example, a r/Conspiracy poster quoted another poster's claim that UK Prime Ministers who resign gets to pick their own successor (entirely untrue).

Definitions were mostly provided from dictionaries, but also from legislation (e.g., rules that defined how an election should be conducted) and the Bible (to define orthodoxy on a subject). Some posts explicitly referenced the source's authority e.g. the r/DebateThis poster who emphasised that their definition of a sphere came from Merriam Webster when debunking flat-earth theory. Subsequently another poster blockquoted the definition noting it was a dictionary not a mathematical reference.

Appeals to authority cut across other evidentiary purposes, as with images. Authorities quoted included individuals (e.g., Lenin) and organisations (e.g., Mayo Clinic), without explicitly articulating their relevant authority in the current discussion. Others were more explicit: one r/ChangeMyView poster blockquoted from George Orwell's biographical 'Homage to Catalonia' to demonstrate Orwell's support for socialism. They appealed to Orwell's authority by claim he had fought for socialism in the Spanish Civil War. They supported their claim by blockquoting Orwell's biographer and providing links. This purpose co-occurred with the others.

4.4.3 Summary and Analysis. Blockquotes more often cited other posters than external sources. The types of external source quoted mirrored some of the types in the 'links' findings; the most common were *academic publications* and *news media articles*. Purposes of using blockquotes to cite evidence mirrored many of the purposes of citing images and links. This included providing *facts*, *numerical data* and *primary evidence*, and *appealing to authority*. 'Good' use of blockquoting (i.e., traceable and authoritative) almost always

occurred in r/DebateThis, while blockquotes in r/Conspiracy and r/ChangeMyView were usually unattributed.

5 DISCUSSION

Our study offers considerable new insight into the role of evidence on Reddit. In the sole two prior studies on online evidentiary practices, both on Reddit, one had found that only 7% of posts used external evidence, though it focused on links only [34] and the other [49] focused only on a single community. We break these insights down into a discussion of social norms and evidentiary practice, and a reflection on how technology might support good evidentiary practice.

5.1 Social Norms and Evidentiary Practice

On Reddit, each subreddit can set its own rules and guidelines for contributions, and these dictate community norms. Previous research has found variances in linguistic toxicity [51] due to community norms, and that norms vary between subreddits [13]. We have shown that evidentiary norms also vary between subreddits. Taking the case of links, the number used in r/Conspiracy and r/ChangeMyView (c. 5% of posts) reflected previous studies of Reddit, while the 33% found in r/DebateThis was much more akin to links used in education focused online communities (44% [39]).

While prior literature has found variances in evidence provision between platforms attributable to social norms [8], we have identified differences within a single platform, i.e. Reddit. This extends beyond just volume and includes the types and purposes of evidence provided.

Our results somewhat support Goodnight's classification of spheres [26]. R/DebateThis, posts have a number of features of *technical evidence*, e.g. quotes are more likely to be from external sources, especially academic papers, and they are more likely to appeal to authority or provide primary evidence than in other subreddits. R/ChangeMyView shows a mixed pattern of evidence, with significantly more links to numerical data and news articles, but also consistent hearsay. In a discussion where only one person is to be convinced, this mixed pattern reflects the needs of the *personal sphere*. Despite the same interface tools being available, evidence provided in r/Conspiracy was least likely to be authoritative e.g. fewer academic reports, and more images. R/Conspiracy evidence is also highly news focused, supporting earlier findings that 'dramatic events' like the Georgia Guidestones incident are key drivers of participation [59]. R/Conspiracy was also most likely to use image-based evidence, (arguably the least reliable format, as images are easily manipulated). When images are screenshots, traceability of the source is reduced, making it harder for readers to assess evidence quality [1]. Nonetheless, evidence was being used, and much of it appealed to authority, meaning evidentiary practice r/Conspiracy is not as far from good evidentiary practice as one might assume. There are also slips in practice across all three subreddits—even on r/DebateThis not all evidence was high quality. This is likely because fact-checking and provenance provision are complex and labour-intensive [14; 33]; reducing this workload would increase good practice in communities where it is normative.

Arguably, as social norms are so influential on Reddit a macro-level shift in social norms toward good evidentiary practice could

entrench good practice across Reddit. Such a shift would require widespread information literacy skills: e.g. analysing sources, and detecting bias in information presentation and in one's own response to it [1; 30; 35]. Inculcating strong critical and evidentiary practices as socially normative early on is most likely to be effective [21]. Focusing only on evidence, though, would privilege technical discussions to the exclusion of other kinds. Even social conversations offer a chance to encounter information that can result in enjoyable experiences [41], new ideas [43], spark creativity [40] or even be view- and life-changing [43]. Not all conversations should be held to technical evidentiary standards, and individual subreddits can (and should) continue to set practice. Where better evidence is called for though, more technical support could be offered.

5.2 Digital Support for Evidentiary Practice

It is notable in our data that blockquoting previous posters, which automatically links to the previous post, results in such quotes always linking to the source quoted. This suggests that there are opportunities for digital support for good evidentiary practice where it is appropriate to the subreddit. Such support might include prompting posters for links to source material for images and blockquotes, resulting in posts that are more traceable, and thus support robust evaluation [1]. Communities may also provide automated links to fact- and bias- checking services where, e.g., news media articles are posted, helping readers to assess evidence, and potentially encouraging posters to provide higher quality evidence.

Supporting posters to question the credibility of evidence, as seen on r/DeBunkThis [28] might promote good practice. One digital example of a lightweight way to question evidence Wikipedia's [citation needed] flag. Lightweight strategies for suggesting a citation is needed—ones that could be completed with a single click—would both encourage the flagging questionable evidence [37], and provide useful information about information credibility to lurkers [3]. Flags could also be used to indicate the use of counterevidence, selective citation, situations where the evidence has changed, or sources of questionable quality. Of course, each of these flags can be weaponised, though moderation on Reddit has been shown to be highly effective in addressing bad behaviour [12; 13; 59].

Alongside opportunities for general support for evidence provision, our findings also highlight the potential to better support providing *specific* types of evidence. For example, when citing external sources, posters might be prompted, when using:

Blockquotes, to provide a link to the *original source*;

Images, to provide a link to the *original source* of: *photographs* (including similar photos that corroborate or question visual proof), *screengrabs* (e.g., of a news article, or social media feed/post), *images that have been annotated post-hoc*, *video stills* (i.e., the video itself) etc. Also, to *flag when the authority appealed to is bogus* (c.f., the fake Georgia Bureau of Investigation logo); OCR of text and links in images would also improve traceability

Links, to provide the *original source of numerical data*, to allow for greater scrutability and to ensure numbers are not being manipulated to tell a biased story. The site being linked to could also be checked for authority, and posters encouraged to use higher authority sites in more technical discussions.

Our typology (of types and purposes of evidence provided in three contrasting Reddit communities) might also be leveraged to assist automatic content classification of online discussion posts. This, in turn, could help online community managers monitor and assess the evidence provision practices of community members, with a view of enhancing them for community (and ideally pro-social) benefit. It could also help researchers and social media advocacy groups characterise different communities with respect to their evidence provision practices, with a view of identifying those with particularly strong or weak practices. Finally, it could support online community users in finding online communities with strong evidentiary practices, with a view of joining reliable, trustworthy communities. Creating technical support to leverage this typology to achieve these goals is an important area for future research.

6 CONCLUSIONS

In this paper, we have presented a mixed method analysis of how and why images, links and blockquotes are used to provide evidence in three very different online communities on Reddit—r/Conspiracy, r/ChangeMyView and r/DeBunkThis. Based on this analysis, we have defined a typology of evidentiary types and purposes. This typology can be adopted and adapted by future researchers to reason about the provision of evidence in other online communities and can be used to support automated detection and processing of public data in online communities.

By comparing the use of evidence between these subreddits, we have demonstrated that community norms influence the kinds of evidence provided, and the way they are provided. While r/ChangeMyView and r/DeBunkThis demonstrate broadly similar patterns of evidence use, r/Conspiracy is more complex. On the one hand, we might expect those participating in a thread on conspiracy theories to eschew evidence altogether, and this is simply not the case. On the other, r/Conspiracy does reflect less evidence use and the evidence that *is* used is often not credible, with plenty of demonstrably false or potentially misleading content cited. However, this is not the full picture; sometimes credible and authoritative sources *are* used, but often highly-selectively in ways that might serve to mislead others. Further research is needed to understand the complex nature of evidence provision in r/Conspiracy. It is clear from this comparison that community norms mediate how evidence is provided and consumed in these online communities.

In a world of abundant misinformation, supporting not just individuals, but entire communities to set sound norms around good evidentiary practice is likely to increase the general pro-sociality of online communities. Knowing that these community norms make a difference is a first important step in this direction. By empowering online communities to embrace strong evidence provision practices, we can ensure greater transparency and scrutability in evidence provision. Although not all communities will want to work to this end, raising the evidence provision bar is likely to make those who do not improve their standards of evidence to stand out, thereby becoming a credibility indicator in its own right. By supporting more and better use of evidence, we can shape online communities to better serve people and society.

REFERENCES

- [1] Bartlett, J.C., 2020. Information Literacy and Science Misinformation. In *Navigating Fake News, Alternative Facts, and Misinformation in a Post-Truth World*. K. Dalkir and R. Katz Eds. IGI Global, Hershey, PA, USA, 1-17. DOI= <http://dx.doi.org/10.4018/978-1-7998-2543-2.ch001>.
- [2] Blakeslee, S., 2004. The CRAAP test. *LOEX Quarterly* 31, 3, 4.
- [3] Bode, L. and Vraga, E.K., 2018. See Something, Say Something: Correction of Global Health Misinformation on Social Media. *Health Communication* 33, 9, 1131-1140. DOI= <http://dx.doi.org/10.1080/10410236.2017.1331312>.
- [4] Boyd, J., 2002. Public and Technical Interdependence: Regulatory Controversy, Out-Law Discourse, and the Messy Case of Olestra. *Argumentation and Advocacy* 39, 2, 91-109. DOI= <http://dx.doi.org/10.1080/00028533.2002.11821579>.
- [5] Bozdag, E. and van den Hoven, J., 2015. Breaking the filter bubble: democracy and design. *Ethics and Information Technology* 17, 4, 249-265. DOI= <http://dx.doi.org/10.1007/s10676-015-9380-y>.
- [6] Bruckman, A., Danis, C., Lampe, C., Sternberg, J., and Waldron, C., 2006. Managing deviant behavior in online communities. In *Proc. CHI EA 06* (Montréal, Québec, Canada), ACM, 1125458, 21-24. DOI= <http://dx.doi.org/10.1145/1125451.1125458>.
- [7] Bruckman, A., Luther, K., and Fiesler, C., 2015. When should we use real names in published accounts of internet research. In *Digital research confidential: The secrets of studying behavior online*, 243-258.
- [8] Bruns, A., 2019. It's not the technology, stupid: How the 'Echo Chamber' and 'Filter Bubble' metaphors have failed us. In *Proc. IAMCR 19* (Madrid, Spain), International Association for Media and Communication Research.
- [9] Carrion, M.L., 2017. "You need to do your research": Vaccines, contestable science, and maternal epistemology. *Public Understanding of Science* 27, 3, 310-324. DOI= <http://dx.doi.org/10.1177/0963662517728024>.
- [10] Chan, M.-p.S., Jones, C.R., Hall Jamieson, K., and Albarracín, D., 2017. Debunking: A Meta-Analysis of the Psychological Efficacy of Messages Countering Misinformation. *Psyc. Sci* 28, 11, 1531-1546. DOI= <http://dx.doi.org/10.1177/0956797617714579>.
- [11] Chancellor, S., Hu, A., and Choudhury, M.D., 2018. Norms Matter: Contrasting Social Support Around Behavior Change in Online Weight Loss Communities. In *Proc. CHI '18* (Montreal QC, Canada), Association for Computing Machinery, Paper 666. DOI= <http://dx.doi.org/10.1145/3173574.3174240>.
- [12] Chandrasekharan, E., Pavalanathan, U., Srinivasan, A., Glynn, A., Eisenstein, J., and Gilbert, E., 2017. You Can't Stay Here: The Efficacy of Reddit's 2015 Ban Examined Through Hate Speech. In *Proc. CSCW 17*, Article 31. DOI= <http://dx.doi.org/10.1145/3134666>.
- [13] Chandrasekharan, E., Samory, M., Jhaver, S., Charvat, H., Bruckman, A., Lampe, C., Eisenstein, J., and Gilbert, E., 2018. The Internet's Hidden Rules: An Empirical Study of Reddit Norm Violations at Micro, Meso, and Macro Scales. In *Proc. CSCW 17*, Article 32. DOI= <http://dx.doi.org/10.1145/3274301>.
- [14] Chen, C.-C. and Roth, C., 2012. {{Citation needed}}: the dynamics of referencing in Wikipedia. In *Proc. WikiSym 12* (Linz, Austria), Association for Computing Machinery, Article 8. DOI= <http://dx.doi.org/10.1145/2462932.2462943>.
- [15] Chen, X., Sin, S.-C.J., Theng, Y.-L., and Lee, C.S., 2015. Why Do Social Media Users Share Misinformation? In *Proc. JCDL 15* (Knoxville, Tennessee, USA), Association for Computing Machinery, 111-114. DOI= <http://dx.doi.org/10.1145/2756406.2756941>.
- [16] Colliander, J., 2019. "This is fake news": Investigating the role of conformity to other users' views when commenting on and spreading disinformation in social media. *Comp. Hum. Behav.* 97, 202-215. DOI= <http://dx.doi.org/10.1016/j.chb.2019.03.032>.
- [17] Connaway, L.S., Julien, H., Seadle, M., and Kasprak, A., 2017. Digital literacy in the era of fake news: Key roles for information professionals. *ASIST Proceedings* 54, 1, 554-555. DOI= <http://dx.doi.org/10.1002/pr2.2017.14505401070>.
- [18] de Souza, C.S. and Preece, J., 2004. A framework for analyzing and understanding online communities. *Interacting with Computers* 16, 3, 579-610. DOI= <http://dx.doi.org/10.1016/j.intcom.2003.12.006>.
- [19] Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H.E., and Quattrociocchi, W., 2016. The spreading of misinformation online. *PNAS* 113, 3, 554. DOI= <http://dx.doi.org/10.1073/pnas.1517441113>.
- [20] Draws, T., Inel, O., Tintarev, N., Baden, C., and Timmermans, B., 2022. Comprehensive Viewpoint Representations for a Deeper Understanding of User Interactions With Debated Topics. In *Proc. CHIIR 22* (Regensburg, Germany), Association for Computing Machinery, 135-145. DOI= <http://dx.doi.org/10.1145/3498366.3505812>.
- [21] Farrell, J., McConnell, K., and Brulle, R., 2019. Evidence-based strategies to combat scientific misinformation. *Nature Climate Change* 9, 3, 191-195. DOI= <http://dx.doi.org/10.1038/s41558-018-0368-6>.
- [22] Flaxman, S., Goel, S., and Rao, J.M., 2016. Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Q* 80, S1, 298-320. DOI= <http://dx.doi.org/10.1093/poq/nfw006>.
- [23] Forte, A., Andalibi, N., Gorichanaz, T., Kim, M.C., Park, T., and Halfaker, A., 2018. Information Fortification: An Online Citation Behavior. In *Proc. GROUP 18* (Sanibel Island, Florida, USA), Association for Computing Machinery, 83-92. DOI= <http://dx.doi.org/10.1145/3148330.3148347>.
- [24] Gilbert, S.A., 2020. "I run the world's largest historical outreach project and it's on a cesspool of a website." Moderating a Public Scholarship Site on Reddit: A Case Study of r/AskHistorians. In *Proc. CSCW 20*, Article 19. DOI= <http://dx.doi.org/10.1145/3392822>.
- [25] Golbeck, J., 2012. The twitter mute button: a web filtering challenge. In *Proc. CHI 2012* (Austin, Texas, USA), Association for Computing Machinery, 2755-2758. DOI= <http://dx.doi.org/10.1145/2207676.2208673>.
- [26] Goodnight, G.T., 1982. The Personal, Technical, and Public Spheres of Argument: A Speculative Inquiry into the Art of Public Deliberation. *J. Amer. Forensic Assn* 18, 4, 214-227. DOI= <http://dx.doi.org/10.1080/00028533.1982.11951221>.
- [27] Halpern, S., 2019. Why The UK Condemned Facebook for Fuelling Fake News. In *New Yorker*.
- [28] He, L. and He, C., 2022. Help Me #DebunkThis: Unpacking Individual and Community's Collaborative Work in Information Credibility Assessment. In *Proc. CSCW 22* (Taipei, Taiwan), ACM. DOI= <http://dx.doi.org/10.1145/3555138>.
- [29] Helberger, N., Karppinen, K., and D'Acunto, L., 2018. Exposure diversity as a design principle for recommender systems. *Inf. Comm & Soc.* 21, 2, 191-207. DOI= <http://dx.doi.org/10.1080/1369118X.2016.1271900>.
- [30] Jones-Jang, S.M., Mortensen, T., and Liu, J., 2021. Does Media Literacy Help Identification of Fake News? Information Literacy Helps, but Other Literacies Don't. *Amer. Behav. Sci* 65, 2, 371-388. DOI= <http://dx.doi.org/10.1177/0002764219869406>.
- [31] Khan, M.L. and Idris, I.K., 2019. Recognise misinformation and verify before sharing: a reasoned action and information literacy perspective. *Behav. & IT* 38, 12, 1194-1212. DOI= <http://dx.doi.org/10.1080/0144929X.2019.1578828>.
- [32] Klein, C., Clutton, P., and Dunn, A.G., 2019. Pathways to conspiracy: The social and linguistic precursors of involvement in Reddit's conspiracy theory forum. *PLOS One* 14, 11, e0225098. DOI= <http://dx.doi.org/10.1371/journal.pone.0225098>.
- [33] Lampe, C., Zube, P., Lee, J., Park, C.H., and Johnston, E., 2014. Crowdsourcing civility: A natural experiment examining the effects of distributed moderation in online forums. *Gov Info Q.* 31, 2, 317-326. DOI= <http://dx.doi.org/https://doi.org/10.1016/j.giq.2013.11.005>.
- [34] Leverich, D., 2018. *Evidence in Online Political Discourse: How Everyday Citizens Argue about Politics on Social Media*. University of Dayton, Dayton, OH.
- [35] Levy, N., 2022. Do your own research! *Synthese* 200, 5, 356. DOI= <http://dx.doi.org/10.1007/s11229-022-03793-w>.
- [36] Lewandowsky, S., Ecker, U.K.H., Seifert, C.M., Schwarz, N., and Cook, J., 2012. Misinformation and Its Correction: Continued Influence and Successful Debiasing. *Psyc Sci in Public Interest* 13, 3, 106-131. DOI= <http://dx.doi.org/10.1177/1529100612451018>.
- [37] Lu, J., Chiu, M.M., and Law, N.W., 2011. Collaborative argumentation and justifications: A statistical discourse analysis of online discussions. *Comput. Hum. Behav.* 27, 2, 946-955. DOI= <http://dx.doi.org/https://doi.org/10.1016/j.chb.2010.11.021>.
- [38] Matias, J.N., 2019. Preventing harassment and increasing group participation through social norms in 2,190 online science discussions. *PNAS* 116, 20, 9785-9789. DOI= <http://dx.doi.org/10.1073/pnas.1813486116>.
- [39] McKay, D., Makri, S., Gutierrez-Lopez, M., MacFarlane, A., Missaoui, S., Porlezza, C., and Cooper, G., 2020. We are the Change that we Seek: Information Interactions During a Change of Viewpoint. In *Proc. CHIIR 20* (Vancouver BC, Canada), Association for Computing Machinery, 173-182. DOI= <http://dx.doi.org/10.1145/3343413.3377975>.
- [40] McKay, D., Owyong, K., Makri, S., and Lopez, M.G., 2022. Turn and Face the Strange: Investigating Filter Bubble Bursting Information Interactions. In *Proc. CHIIR 22* (Regensburg, Germany), Association for Computing Machinery, 233-242. DOI= <http://dx.doi.org/10.1145/3498366.3505822>.
- [41] Medvedev, A.N., Lambiotte, R., and Delvenne, J.-C., 2019. The Anatomy of Reddit: An Overview of Academic Research. In *Proc. DOOCN 17* (Indianapolis, IN), Springer International Publishing, 183-204. DOI= http://dx.doi.org/10.1007/978-3-030-14683-2_9.
- [42] Mercier, H. and Sperber, D., 2011. Why do humans reason? Arguments for an argumentative theory. *Behav. Brain Sci* 34, 2, 57-74. DOI= <http://dx.doi.org/https://doi.org/10.1017/S0140525X10000968>.
- [43] Nonnecke, B. and Preece, J., 2000. Lurker demographics: counting the silent. In *Proc. CHI 00* (The Hague, The Netherlands), Association for Computing Machinery, 73-80. DOI= <http://dx.doi.org/10.1145/332040.332409>.
- [44] Öcal, A., Xiao, L., and Park, J., 2021. Reasoning in social media: insights from Reddit "Change My View" submissions. *Online Info. Rev.* 45, 7, 1208-1226. DOI= <http://dx.doi.org/10.1108/OIR-08-2020-0330>.
- [45] Oxford English Dictionary, "evidence, n.". Oxford University Press.
- [46] Petty, E., S.C., W., and Tormala, Z., 2003. Persuasion and Attitude Change. In *Handbook of Psychology* Wiley, 353-382. DOI= <http://dx.doi.org/10.1002/0471264385.wei0515>.
- [47] Preece, J., 2000. *Online communities: Designing usability and supporting socialbilty*. John Wiley & Sons, Inc.
- [48] Preece, J., 2001. Online Communities: Usability, Sociabilty, Theory and Methods. In *Frontiers of Human-Centered Computing, Online Communities and Virtual Environments*, R.A. Earnshaw, R.A. Guedj, A.v. Dam and J.A. Vince Eds. Springer London, London, 263-277. DOI= http://dx.doi.org/10.1007/978-1-4471-0259-5_18.

- [49] Priniski, J.H. and Horne, Z., 2018. Attitude Change on Reddit's Change My View. In *Proc. CogSci18* (Madison, WS), Cognitive Science Society.
- [50] Proferes, N., Jones, N., Gilbert, S., Fiesler, C., and Zimmer, M., 2021. Studying Reddit: A Systematic Overview of Disciplines, Approaches, Methods, and Ethics. *Social Media + Society* 7, 2, 20563051211019004. DOI= <http://dx.doi.org/10.1177/20563051211019004>.
- [51] Rajadesingan, A., Resnick, P., and Budak, C., 2020. Quick, Community-Specific Learning: How Distinctive Toxicity Norms Are Maintained in Political Subreddits. In *Proc. ICWSM '20*, 557-568. DOI= <http://dx.doi.org/10.1609/icwsm.v14i1.7323>.
- [52] Rieger, A., Draws, T., Theune, M., and Tintarev, N., 2021. This Item Might Reinforce Your Opinion: Obfuscation and Labeling of Search Results to Mitigate Confirmation Bias. In *Proc. HT21* (Virtual Event, USA), Association for Computing Machinery, 189–199. DOI= <http://dx.doi.org/10.1145/3465336.3475101>.
- [53] Rubin Victoria, L., 2019. Disinformation and misinformation triangle: A conceptual model for “fake news” epidemic, causal factors and interventions. *J Doc* 75, 5, 1013-1034. DOI= <http://dx.doi.org/10.1108/JD-12-2018-0209>.
- [54] Saling, L.L., Mallal, D., Scholer, F., Skelton, R., and Spina, D., 2021. No one is immune to misinformation: An investigation of misinformation sharing by subscribers to a fact-checking newsletter. *PLOS ONE* 16, 8, e0255702. DOI= <http://dx.doi.org/10.1371/journal.pone.0255702>.
- [55] Samory, M. and Mitra, T., 2018. Conspiracies Online: User Discussions in a Conspiracy Community Following Dramatic Events. In *Proc. AAAI 18*.
- [56] Scolyer-Gray, P., 2022. Artistic Works of Fiction and Falsehood: An Analysis of the Production and Consumption of Knowledge on 4chan. Patrick Scolyer-Gray.
- [57] Seargeant, P. and Tagg, C., 2019. Social media and the future of open debate: A user-oriented approach to Facebook's filter bubble conundrum. *Doiscourse Context and Media* 27, 41-48. DOI= <http://dx.doi.org/10.1016/j.dcm.2018.03.005>.
- [58] Snavey, L. and Cooper, N., 1997. The information literacy debate. *J. Acad. Libr.* 23, 1, 9-14. DOI= [http://dx.doi.org/https://doi.org/10.1016/S0099-1333\(97\)90066-5](http://dx.doi.org/https://doi.org/10.1016/S0099-1333(97)90066-5).
- [59] Srinivasan, K.B., Danescu-Niculescu-Mizil, C., Lee, L., and Tan, C., 2019. Content Removal as a Moderation Strategy: Compliance and Other Outcomes in the ChangeMyView Community. In *Proc. CSCW 19*, Article 163. DOI= <http://dx.doi.org/10.1145/3359265>.
- [60] Stefanone, M.A., Vollmer, M., and Covert, J.M., 2019. In News We Trust? Examining Credibility and Sharing Behaviors of Fake News. In *Proc. SMS19* (Toronto, ON, Canada), Association for Computing Machinery, 136–147. DOI= <http://dx.doi.org/10.1145/3328529.3328554>.
- [61] Suh, A. and Wagner, C., 2013. Factors Affecting Individual Flaming in Virtual Communities. In *Proc. HICSS '13*, 3282-3291. DOI= <http://dx.doi.org/10.1109/HICSS.2013.230>.
- [62] Tambuscio, M., Ruffo, G., Flammini, A., and Menczer, F., 2015. Fact-checking Effect on Viral Hoaxes: A Model of Misinformation Spread in Social Networks. In *Proc. WWW 15* (Florence, Italy), Association for Computing Machinery, 977–982. DOI= <http://dx.doi.org/10.1145/2740908.2742572>.
- [63] Terveen, L., Konstan, J., and Lampe, C., 2014. Study, Build, Repeat: Using Online Communities as a Research Platform. In *Ways of Knowing in HCI*, J.S. Olson and W.A. Kellogg Eds. Springer New York, 95-117. DOI= http://dx.doi.org/10.1007/978-1-4939-0378-8_5.
- [64] Tewell, E., 2015. A decade of critical information literacy: A review of the literature. *Comm. Info. Lit.* 9, 1, 2. DOI= <http://dx.doi.org/10.15760/comminfolit.2015.9.1.174>.
- [65] Vranic, A., Hromatko, I., and Tonković, M., 2022. "I Did My Own Research": Overconfidence, (Dis)trust in Science, and Endorsement of Conspiracy Theories. *Front. Psychol.* 13, 931865. DOI= <http://dx.doi.org/10.3389/fpsyg.2022.931865>.
- [66] Weeks, B.E., Ksiazek, T.B., and Holbert, R.L., 2016. Partisan Enclaves or Shared Media Experiences? A Network Approach to Understanding Citizens' Political News Environments. *J Broadcasting Elec Media* 60, 2, 248-268. DOI= <http://dx.doi.org/10.1080/08838151.2016.1164170>.
- [67] Wu, L., Morstatter, F., Carley, K.M., and Liu, H., 2019. Misinformation in Social Media: Definition, Manipulation, and Detection. *SIGKDD Explor. Newsl.* 21, 2, 80–90. DOI= <http://dx.doi.org/10.1145/3373464.3373475>.
- [68] Xiao, L. and Khazaei, T., 2019. Changing Others' Beliefs Online: Online Comments' Persuasiveness. In *Proc. SMS18* (Toronto, ON, Canada), Association for Computing Machinery, 92–101. DOI= <http://dx.doi.org/10.1145/3328529.3328549>.
- [69] Yom-Tov, E., Dumais, S., and Guo, Q., 2013. Promoting Civil Discourse Through Search Engine Diversity. *Soc Sci Comp Rev* 32, 2, 145-154. DOI= <http://dx.doi.org/10.1177/0894439313506838>.
- [70] Zannettou, S., Sirivianos, M., Blackburn, J., and Kourtellis, N., 2019. The Web of False Information: Rumors, Fake News, Hoaxes, Clickbait, and Various Other Shenanigans. *J. Data and Inf. Quality* 11, 3, Article 10. DOI= <http://dx.doi.org/10.1145/3309699>.
- [71] Zhang, H., Alim, M.A., Li, X., Thai, M.T., and Nguyen, H.T., 2016. Misinformation in Online Social Networks: Detect Them All with a Limited Budget. *ToIS* 34, 3, Article 18. DOI= <http://dx.doi.org/10.1145/2885494>.
- [72] Zhou, Y. and Farzan, R., 2021. Designing to Stop Live Streaming Cyberbullying: A case study of Twitch Live Streaming Platform. In *Proc. C&T 21* (Seattle, WA, USA), Association for Computing Machinery, 138–150. DOI= <http://dx.doi.org/10.1145/3461564.3461574>.