

Trust and Distrust in Online Fact-Checking Services

By **Petter Bae Brandtzaeg and Asbjørn Følstad**

While the Internet has the potential to give people ready access to relevant and factual information, social media sites like Facebook and Twitter have made filtering and assessing online content increasingly difficult due to its rapid flow and enormous volume. In fact, 49% of social media users in the U.S. in 2012 received false breaking news through social media.[8] Likewise, a survey by Silverman[11] suggested that in 2015 false rumors and misinformation disseminated further and faster than ever before due to social media. Political analysts continue to discuss the growing misinformation and fake news in social media and its effect on the 2016 U.S. presidential election.

Such misinformation challenges the credibility of the Internet as a venue for public information and debate. In response over the past five years, a proliferation of outlets has provided fact checking and debunking of online content. Fact-checking services, say Kriplean et al.,[6] provide "... evaluation of verifiable claims made in public statements through investigation of primary and secondary sources." An international census from 2017 counted 114 active fact-checking services, a 19% increase from the previous year.[12] To benefit from this trend, Google News in 2016 let news providers tag news articles or their content with fact-checking information "... to help readers find fact checking in large news stories." [3] Any organization may use the fact-checking tag, if it is non-partisan, transparent, and targets a range of claims within an area of interest and not just one single person or entity.

However, research in fact checking has scarcely paid attention to the general public's view of fact-checking services, focusing instead on how people's beliefs and attitudes change in response to facts that contradict their own pre-existing opinions. This research suggests fact checking in general may be unsuccessful at reducing misperceptions, especially among the people most prone to believe them.[9] People often ignore facts that contradict their current beliefs,[2,13] particularly in politics and controversial social issues.[9] Consequently, the more political or controversial issues a fact-checking service covers, the more it needs to build a reputation for usefulness and trustworthiness.

Research suggests the trustworthiness of fact-checking services depends on their origin and ownership, which may affect integrity perceptions[10] and the transparency of their fact-checking process.[4] Despite these observations, we are unaware of any research that has examined users' perceptions of fact-checking services. Addressing the gap in current knowledge, we investigated this research question: How do social media users perceive the trustworthiness and usefulness of fact-checking services?

Fact-checking services differ in terms of their organizational aim and funding,[10] as well as their areas of concern,[11] that in turn may affect their trustworthiness. As outlined in Figure 1, the universe of fact-checking services can be divided into three general categories based on their areas of concern: political and public statements in general, corresponding to the fact checking of politicians, as discussed by Nyhan and Reifler;[9] online rumors and hoaxes, reflecting the need for debunking services, as discussed by Silverman;[11] and specific topics or controversies or particular conflicts or narrowly scoped issues or events (such as the ongoing Ukraine conflict).

In our study, we focused on three services—Snopes, FactCheck.org, and StopFake—all included in

the Duke Reporters' Lab's online overview of fact checkers (<http://reporterslab.org/fact-checking/>). They represent three categories of fact checkers, from online rumors to politics to a particular topic of controversy, as in Figure 1, and differences in organization and funding. As a measure of their popularity, as of June 20, 2017, Snopes had 561,650 likes on Facebook, FactCheck.org 806,814, and StopFake 52,537.

We study Snopes because of its aim to debunk online rumors, fitting the first category in Figure 1. This aim is shared by other such services, including HoaxBusters and the Swedish service Viralgranskaren. Snopes is managed by a small volunteer organization emerging from a single-person initiative and funded through advertising revenue.

We study FactCheck.org because it monitors the factual accuracy of what is said by major figures in politics. Other similar services include PolitiFact (U.S.) and Full Fact (U.K.) listed in the second category in Figure 1. FactCheck.org is a project of the Annenberg Public Policy Center of the Annenberg School for Communication at the University of Pennsylvania in U.S. FactCheck.org is supported by university funding and individual donors and has been a source of inspiration for other fact-checking projects.

We study StopFake because it addresses a highly specific topic—the ongoing Ukraine conflict. It thus resembles other highly focused fact-checking initiatives (such as #Refugeecheck, an initiative to fact check reports on the refugee crises in Europe). StopFake is an initiative by the Kyiv Mohyla Journalism School in Kiev, Ukraine, and thus a European-based service. Snopes and FactCheck.org are U.S. based, like more than a third of the fact-checking services identified by Duke Reporters' Lab.[12]

All three provide their fact checking through their own websites, as well as through Facebook and Twitter. Figure 2 is an example of a Twitter post with content checked by Snopes.

Analyzing Social Media Conversations

To explore how social media users perceive the trustworthiness and usefulness of fact-checking services, we applied a research approach designed to take advantage of unstructured social media conversations (see Figure 3).

While investigations of trust and usefulness often rely on structured data from questionnaire-based surveys, social media conversations represent a highly relevant data source for our purpose, as they arguably reflect the raw, authentic perceptions of social media users. Xu et al.[16] claim it is beneficial to listen to, analyze, and understand citizens' opinions through social media to improve societal decision-making processes and solutions. For example, Xu et al.[16] wrote, "social media analytics has been applied to explain, detect, and predict disease outbreaks, election results, macroeconomic processes (such as crime detection), (...) and financial markets (such as stock price)."[16] Social media conversations take place in the everyday context of users likely to be engaged in fact-checking services. This approach may provide a more unbiased view of people's perceptions than, say, a questionnaire-based approach. The benefit of gathering data from users in their specific social media context does not imply that our data is representative. Our data lacks important information about-user demographics, limiting our ability to claim generality for the entire user population. Despite this potential drawback, our data does offer new insight into how social media users view the usefulness and trustworthiness of various categories of fact-checking services.

For data collection, we used Meltwater Buzz, an established service for social media monitoring.

This service crawled data from social media conversations in blogs, discussion forums, online newspaper discussion threads, Twitter, and Facebook. Meltwater Buzz is designed to crawl all blogs (such as <https://wordpress.com/>), discussion forums (such as <https://offtopic.com/>), and online newspapers (such as <https://www.washingtonpost.com/>) requested by Meltwater customers, thus representing a convenience sample, albeit large. Meltwater Buzz collects various amounts of data from each platform; for example, it crawls all posts on Twitter but only the Facebook pages with 3,500 likes or groups of more than 500 members. This limitation in Facebook data partly explains why the overall number of posts we collected—1,741—was not very great.

To collect opinions about social media user perceptions of Snopes and FactCheck.org, we applied the search term “[service name] is,” or “Snopes is,” “FactCheck.org is,” and “FactCheck is.” This term was intended to reflect how people start a sentence when formulating opinions. StopFake is a relatively less-known service. We thus selected a broader search string—“StopFake”—to be able to collect enough relevant opinions. The searches returned a data corpus of 1,741 posts going back six months in time—October 2014 to March 2015—as in Figure 3. By “posts,” we mean written contributions by individual users. To create a sufficient dataset for analysis, we removed all duplicates, including a small number of non-relevant posts lacking personal opinions about fact checkers. This filtering process resulted in a dataset of 595 posts.

We then performed content analysis, coding all posts to identify and investigate patterns within the data,[1] to reveal the perceptions users express in social media about the three fact-checking services we investigated. We analyzed their perceptions of the usefulness of fact-checking services through a usefulness construct similar to the one used by Tsakonas et al.[14] Hence, “usefulness” concerns the extent the service is perceived as beneficial when doing a specific fact-checking task, often illustrated by positive recommendations and characterizations (such as that the service is

“good” or “great”). Following Mayer et al.’s theoretical framework,[7] we categorized trustworthiness according to the perceived ability, benevolence, and integrity of the fact-checking services. “Ability” concerns the extent the service is perceived as having available the needed skills and expertise, as well as being reputable and well regarded. “Benevolence” refers to the extent the service is perceived as intending to do good, beyond what would be expected from an egocentric motive. “Integrity” targets the extent the service is generally viewed as adhering to an acceptable set of principles, in particular being independent, unbiased, and fair.

Since we found posts typically reflect rather polarized perceptions of the studied fact-checking services, we also grouped the codes manually according to sentiment, whether positive or negative. However, some posts describe the services in a plain and objective manner. We thus coded them using a positive sentiment (see Table 1) because they refer to the service as a source for fact checking, and users are likely to reference fact-checking sites because they see them as useful.

For reliability, both researchers in this study did the coding. One coded all the posts, and the second then went through all the assigned codes, a process that was repeated twice. Finally, both researchers went through all comments for which an alternative code had been suggested to decide on the final coding, a process that recommended an alternative coding for 153 posts (26%).

A post could include more than one of the analytical themes, so 30% of the posts were thus coded as addressing two or more themes.

Results

Despite the potential benefits of fact-checking services, Figure 4 shows that the majority of the posts on the two U.S.-based services expressed negative sentiment, with Snopes at 68% and

FactCheck.org at 58%. Most posts on the Ukraine-based StopFake (78%) reflected positive sentiment.

The stated reasons for negative sentiment typically concerned one or more of the trustworthiness themes rather than usefulness. For example, for Snopes and FactCheck.org, the negative posts often expressed concern over lack in integrity due to perceived bias toward the political left. Negative sentiment pertaining to the ability and benevolence of the services were also common. The few critical comments on usefulness were typically aimed at discrediting a service, as by, say, characterizing it as “satirical” or “a joke.”

Positive posts were more often related to usefulness. For example, the stated reasons for positive sentiment toward StopFake typically concerned the service’s usefulness in countering pro-Russian propaganda and trolling and in the information war associated with the ongoing Ukraine conflict.

In line with a general notion of an increasing need to interpret and act on information and misinformation in social media,[6,11] some users in our study discussed fact-checking sites as important elements of an information war.

Snopes. The examples in Table 2 reflect how negative sentiment in the posts we analyzed on Snopes was rooted in issues pertaining to trustworthiness. Integrity issues typically involved a perceived “left-leaning” political bias in the people behind the service. Pertaining to benevolence, users in our study said Snopes is part of a larger left-leaning, or “liberal,” conspiracy often claimed to be funded by George Soros, whereas comments on ability typically targeted lack of expertise in the people running the service. Some negative comments on trustworthiness may be seen as a rhetorical means of discrediting the service. Posts expressing positive sentiment mainly argued for

the usefulness of the service, claiming that Snopes is a useful resource for checking up on the veracity of Internet rumors.

FactCheck.org. The patterns in the posts we analyzed for FactCheck.org resemble those for Snopes. As in Table 3, the most frequently mentioned trustworthiness concerns related to service integrity; as for Snopes, users said the service was politically biased toward the left. Posts concerning benevolence and ability were also relatively frequent, reflecting user concern regarding the service as a contributor to propaganda or doubts about its fact-checking practices.

StopFake. As in Table 4, the results for StopFake show more posts expressing positive sentiment than we found for Snopes and FactCheck.org. In particular, the posts in the study pointed out that StopFake helps debunk rumors seen as Russian propaganda in the Ukraine conflict.

Nevertheless, the general pattern in the reasons users gave us for positive and negative sentiment for Snopes and FactCheck.org also held for StopFake. The positive posts were typically motivated by usefulness, whereas the negative posts reflected the sentiment that StopFake is politically biased (“integrity”), a “fraud,” a “hoax,” or part of Ukraine propaganda machinery (“benevolence”).

Discussion

We found users in the study with positive perceptions typically extol the usefulness of fact-checking services, whereas users with negative opinions cited trust issues. This pattern emerged across the three different services. In the following sections, we discuss how these findings provide new insight into trustworthiness as a key challenge when countering online rumors and misinformation[2,9] and why ill-founded beliefs may have such online reach, even though the beliefs are corrected by such prominent fact checkers as Snopes, FactCheck.org, and StopFake.

Usefulness. Users in our sample with a positive view of the services mainly pointed to their usefulness. While everyone should exercise caution when comparing the different fact checkers, topic-specific StopFake is perceived as more useful than Snopes and FactCheck.org. One reason might be that a service targeting a specific topic faces less criticism because it attracts a particular audience that seeks facts supporting their own view. For example, StopFake users target anti-Russian and pro-Ukrainian readers. Another, more general, reason might be positive perceptions are motivated by user needs pertaining to a perceived high load of misinformation, as in the case of the Ukraine conflict, where media reports and social media are seen as overflowing with propaganda. Others highlighted the general ease information may be filtered or separated from misinformation through sites like Snopes and FactCheck.org, as expressed like this:

“As you pointed out, it doesn’t take that much effort to see if something on the Internet is legit, and Snopes is a great place to start. So why not take that few seconds of extra effort to do that, rather than creating and sharing misleading items.”

This finding suggests there is an increasing demand for fact-checking services,[6] while at the same time a substantial proportion of social media users who would benefit from fact-checking services do not use them sufficiently. Such services should thus be even more active on social media sites like Facebook and Twitter, as well as in online discussion forums, where easy access to fact checking is needed.

Trustworthiness. Negative perceptions and opinions about fact-checking services seem to be motivated by basic distrust rather than by rational argument. For some users in our sample, lack of trust extends beyond a particular service to encompass the entire social and political system. Users

with negative perceptions thus seem trapped in a perpetual state of informational disbelief.

While one's initial response to statements reflecting a state of informational disbelief may be to dismiss them as the uninformed paranoia of a minority of the public, the statements should instead be viewed as a source of user insight. The reason fact-checking services are often unsuccessful in reducing ill-founded perceptions[9] and people tend to disregard fact checking that goes against their pre-existing beliefs[2,13] may be a lack of basic trust rather than a lack of fact-based arguments provided by the services.

We found such distrust is often highly emotional. In line with Silverman,[11] fact-checking sites must be able to recognize how debunking and fact checking evoke emotion in their users. Hence, the services may benefit from rethinking the way they design and present themselves to strengthen trust among users in a general state of informational disbelief. Moreover, users of online fact-checking sites need to compensate for the lack of physical evidence online by, say, being demonstrably independent, impartial, and able to clearly distinguish fact from opinion.

Rogerson[10] wrote that fact-checking sites exhibit varying levels of rigor and effectiveness. The fact-checking process and even what are considered "facts" may in some cases involve subjective interpretation, especially when actors with partial ties aim to provide fact-checking services. For example, in the 2016 U.S. presidential campaign, the organization "Donald J. Trump for President" invited Trump supporters to join a fact-check initiative, similar to the category "topics or controversies," urging "fact checking" the presidential debates on social media. However, the initiative was criticized as mainly promoting Trump's views and candidacy.[5]

Important questions users of fact-checking sites ask include: Who does the fact checking and how do they do it? What organizations are behind the process? And how does the nature of the

organization influence the results of the fact checking? Fact-checking sites must thus explicate the nuanced, detailed process leading to the presented result while keeping it simple enough to be understandable and useful.[11]

Need for transparency. While fact-checker trustworthiness is critical, fact checkers represent but one set of voices in the information landscape and cannot be expected to be benevolent and unbiased just by virtue of their checking facts. Rather, fact-checking services, more than other types of information service must strive for transparency in their working process, as well as in their origins, organization, and funding sources.

To increase transparency in fact-checking processes, a service might try to take a more horizontal, collaborative approach than is typically seen in today's generation of services. Following Hermida's recommendation[4] for social media journalists, fact checkers could be set up as platforms for collaborative verification and fact checking, relying less on centralized expertise. Forming an interactive relationship with users may also help build trust.[6,7]

Conclusion

We identified a lack of perceived trustworthiness and a state of informational disbelief as potential obstacles to fact-checking services reaching social media users most critical to such services. Table 5 summarizes our overall findings and discussions, outlining related key challenges and our recommendations for how to address them.

Given the exploratory nature of the study, we cannot conclude that our findings are valid for all fact-checking services. In addition, more research is needed to make definite claims on systematic differences between various forms of fact checkers based on their "areas of concern." Nevertheless,

the consistent pattern in opinions we found across three prominent fact-checking services suggests challenges and recommendations that can provide useful guidance for future development in this important area.

Acknowledgments

This work was supported by the European Commission co-funded FP 7 project REVEAL (Project No. FP7-610928, <http://www.revealproject.eu/>) but does not necessarily represent the views of the European Commission. We also thank Marika Lüders of the University of Oslo and three anonymous reviewers for their insightful comments.

References

1. Ezzy, D. *Qualitative Analysis*. Routledge, London, U.K., 2013.
2. Friesen, J.P., Campbell, T.H., and Kay, A.C. The psychological advantage of unfalsifiability: The appeal of untestable religious and political ideologies. *Journal of Personality and Social Psychology* 108, 3 (Nov. 2014), 515–529.
3. Gingras, R. Labeling fact-check articles in Google News. *Journalism & News* (Oct. 13, 2016); <https://blog.google/topics/journalism-news/labeling-fact-check-articles-google-news/>
4. Hermida, A. Tweets and truth: Journalism as a discipline of collaborative verification. *Journalism Practice* 6, 5-6 (Mar. 2012), 659–668.
5. Jamieson, A. ‘Big League Truth Team’ pushes Trump’s talking points on social media. *The Guardian* (Oct. 10, 2016); <https://www.theguardian.com/us-news/2016/oct/10/donald-trump-big-league-truth-team-social-media-debate>
6. Kriplean, T., Bonnar, C., Borning, A., Kinney, B., and Gill, B. Integrating on-demand fact-checking with public dialogue. In *Proceedings of the 17th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (Baltimore, MD, Feb. 15–19). ACM Press, New York, 2014, 1188–1199.
7. Mayer, R.C., Davis, J.H., and Schoorman, F.D. An integrative model of organizational trust.

Academy of Management Review 20, 3 (1995), 709–734.

8. Morejon, R. How social media is replacing traditional journalism as a news source. *Social Media Today Report* (June 28, 2012); <http://www.socialmediatoday.com/content/how-social-media-replacing-traditional-journalism-news-source-infographic>

9. Nyhan, B. and Reifler, J. When corrections fail: The persistence of political misperceptions. *Political Behavior* 32, 2 (June 2010), 303–330.

10. Rogerson, K.S. Fact-checking the fact checkers: Verification web sites, partisanship and sourcing. In *Proceedings of the American Political Science Association* (Chicago, IL, Aug. 29–Sept. 1). American Political Science Association, Washington, D.C., 2013.

11. Silverman, C. *Lies, Damn Lies, and Viral Content. How News Websites Spread (and Debunk) Online Rumors, Unverified Claims, and Misinformation*. Tow Center for Digital Journalism, Columbia Journalism School, New York, 2015; http://towcenter.org/wp-content/uploads/2015/02/LiesDamnLies_Silverman_TowCenter.pdf

12. Stencel, M. International fact-checking gains ground, Duke census finds. Duke Reporters' Lab, Duke University, Durham, NC, Feb. 28, 2017; <https://reporterslab.org/international-fact-checking-gains-ground/>

13. Stroud, N.J. Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior* 30, 3 (Sept. 2008), 341–366.

14. Tsakonas, G. and Papatheodorou, C. Exploring usefulness and usability in the evaluation of open access digital libraries. *Information Processing & Management* 44, 3 (May 2008), 1234–1250.

15. Van Mol, C. Improving web survey efficiency: The impact of an extra reminder and reminder content on web survey response. *International Journal of Social Research Methodology* 20, 4 (May 2017), 317–327.

16. Xu, C., Yu, Y., and Hoi, C.K. Hidden in-game intelligence in NBA players' tweets. *Commun. ACM* 58, 11 (Nov. 2015), 80–89.

Petter Bae Brandtzaeg (pbb@sintef.no) is a senior research scientist at SINTEF in Oslo, Norway.

Asbjørn Følstad (asf@sintef.no) is a senior research scientist at SINTEF in Oslo, Norway

Fact-checking services – areas of concern



Figure 1. Categorization of fact-checking services based on areas of concern.



Figure 2. Example of Snopes debunking a social media rumor on Twitter (March 6, 2016);

<https://twitter.com/snopes/status/706545708233396225>

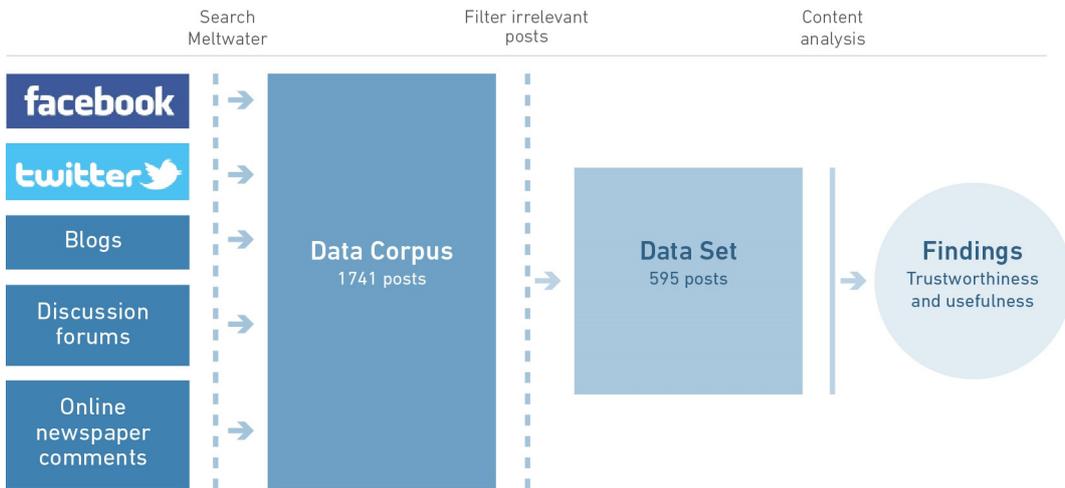


Figure 3. Outline of our research approach; posts collected October 2014 to March 2015.

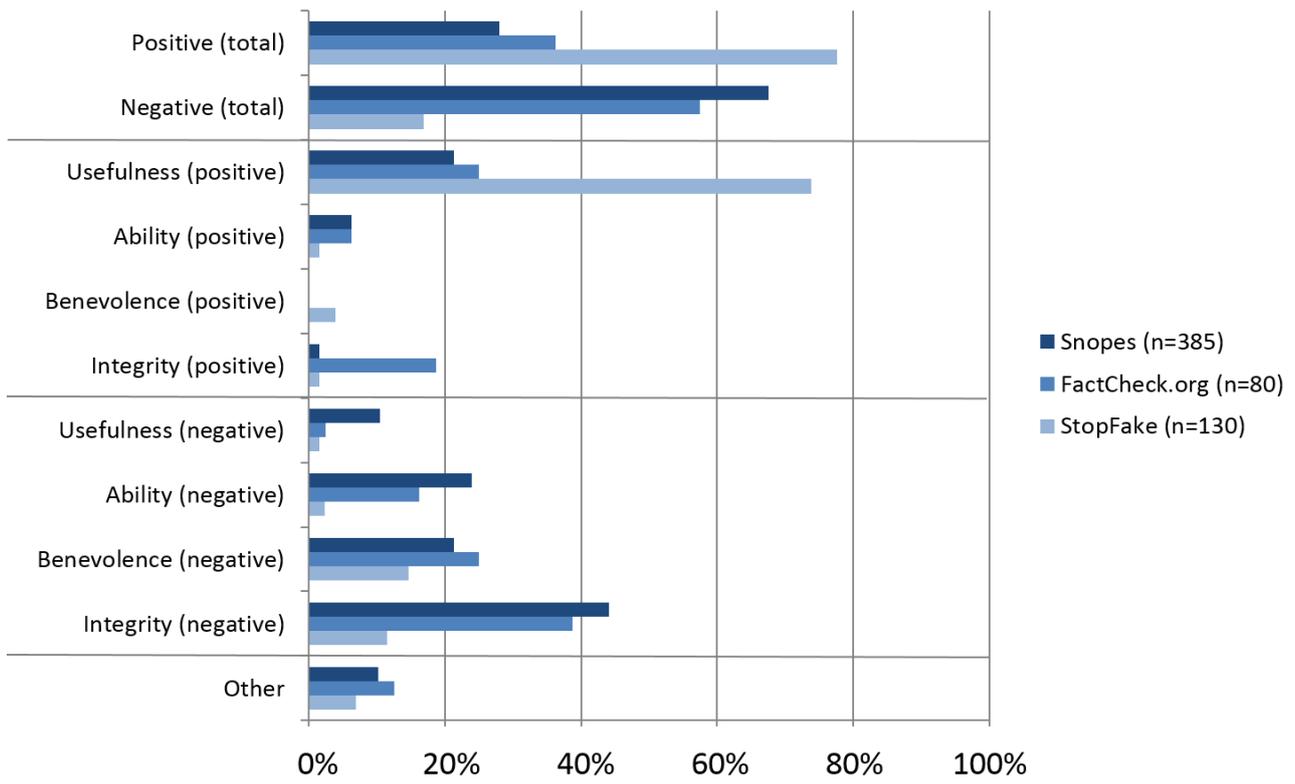


Figure 4. Positive and negative posts related to trustworthiness and usefulness per fact-checking service (in %); “other” refers to posts not relevant for the researching categories (N = 595 posts).

Theme	Sentiment	The service described as
Usefulness	Positive	Useful, serving the purpose of fact-checking.
	Negative	Not as useful, often in a derogatory manner.
Ability	Positive	Reputable, expert, or acclaimed.
	Negative	Lacking in expertise or credibility.
Benevolence	Positive	Aiming for a greater (social) good.
	Negative	Suspected of (social) ill-doing (such as through conspiracy, propaganda, or fraud)
Integrity	Positive	Independent or impartial.
	Negative	Dependent, partial, or politically biased.

Table 1. Coding scheme we used to analyze the data.

Theme	Sentiment	Example
Usefulness	Positive (21%)	Snopes is a wonderful Website for verifying things you see online; it is at least a starting point for research.
	Negative (10%)	Snopes is a joke. Look at its Boston bombing debunking failing to debunking the worst hoax ever ...
Ability	Positive (6%)	[...] Snopes is a respectable source for debunking wives tails, urban legends, even medical myths ...
	Negative (24%)	Heh... Snopes is a man and a woman with no investigative background or credentials who form their opinions solely on internet research; they don't interview anyone. [...]
Benevolence	Positive (0%)	No posts
	Negative (21%)	You are showing your Ignorance by using Snopes ... SNOPEs is a NWO Disinformation System designed to fool the Masses ... SORRY. I Believe NOTHING from Snopes. Snopes is a Disinformation vehicle of the Elitist NWO Globalists. Believe NOTHING from them... [...]
Integrity	Positive (2%)	Snopes is a standard, rather dull fact-checking site, nailing right and left equally. [...]
	Negative (44%)	Snopes is a leftist outlet supported with money from George Soros. Whatever Snopes says I take with a grain of salt ...

Table 2. Snopes and themes we analyzed (n = 385).

Theme	Sentiment	Example
Usefulness	Positive (25%)	[...] You obviously haven't listened to what they say. Also, I hate liars. Fact check is a great tool.
	Negative (3%)	Anyway, "FactCheck" is a joke [...]
Ability	Positive (6%)	The media sources I use must pass a high credibility bar. FactCheck.org is just one of the resources I use to validate what I read....
	Negative (16%)	[...] FactCheck is NOT a confidence builder; see their rider and sources, Huffpo articles ... REALLY?
Benevolence	Positive (0%)	No posts.
	Negative (25%)	FactCheck studies the factual correctness of what major players in U.S. politics say in TV commercials, debates, talks, interviews, and news presentations, then tries to present a best possible fictional and propagand-like version for its target [...]
Integrity	Positive (19%)	When you don't like the message, blame the messenger. FactCheck is non-partisan. It's just that conservatives either lie or are mistaken more ...
	Negative (39%)	FactCheck is a left-leaning opinion. It doesn't check facts ...

Table 3. FactCheck.org and themes we analyzed (n = 80).

Theme	Sentiment	Example
Usefulness	Positive (72%)	Don't forget a strategic weapon of the Kremlin is the "web of lies" spread by its propaganda machine; see antidote http://www.stopfake.org/en/news
	Negative (2%)	[...] Stopfake! HaHaHa. You won, I give up. Next time I will quote 'Saturday Night Live'; there is more truth:))...
Ability	Positive (2%)	[...] by the way the website StopFake.org is a very objective and accurate source exposing Russian propaganda and disinformation techniques. [...]*
	Negative (2%)	[...] Ha Ha ... a flow of lies is constantly sent out from the Kremlin. Really. If so, StopFake needs updates every hour, but the best way it can do that is to find low-grade blog content and make it appear as if it was produced by Russian media [...]
Benevolence	Positive (4%)	[...] Stopfake is devoted to exposing Russian propaganda against the Ukraine. [...]
	Negative (14%)	So now you acknowledge Stopfake is part of Kiev's propaganda. I guess that answers my question [...]
Integrity	Positive (2%)	[...] by the way the website StopFake.org is a very objective and accurate source exposing Russian propaganda and disinformation techniques. [...]
	Negative (11%)	[...] Why should I give any credence to StopFake.org? Does it ever criticize the Kiev regime, in favor of the Donbass position? [...]

Table 4. StopFake and themes we analyzed (n = 130); note * also coded as integrity/positive.

		Challenges	Recommendations
Usefulness		Unrealized potential in public use of fact-checking services	Increase presence in social media and discussion forums
Trustworthiness	Ability	Critique of expertise and reputation	Provide nuanced but simple overview of the fact-checking process where relevant sources are included
	Benevolence	Suspicion of conspiracy and propaganda	Establish open policy on fact checking and open spaces for collaboration on fact checking
	Integrity	Perception of bias and partiality	Ensure transparency on organization and funding. And demonstrable impartiality in fact-checking process

Table 5. Challenges and our related recommendations for fact-checking services.