Assessing Credibility Factors of Short-Form Social Media Posts: A Crowdsourced Online Experiment

Junhao Li
junhao.li@oulu.fi
University of Oulu
Oulu, Finland

Miikka Kuutila
miikka.kuutila@oulu.fi
University of Oulu
Oulu, Finland

Eetu Huusko
eetu.huusko@oulu.fi
University of Oulu
Oulu, Finland

Nimantha Kariyakarawana
vishva.kariyakarawana@oulu.fi
University of Oulu
Oulu, Finland

Marko Savic
marko.savic@oulu.fi
University of Oulu
Oulu, Finland

Nazanin Nakhaie Ahooie
nazanin.nakhaieahooie@oulu.fi
University of Oulu
Oulu, Finland

Simo Hosio
simo.hosio@oulu.fi
University of Oulu
Oulu, Finland

Mika Mäntylä
mika.mantyla@oulu.fi
University of Oulu
Oulu, Finland

ABSTRACT

People commonly turn to the Internet and social media for their information needs. Most popular social media platforms focus on short-form content that can be consumed rapidly. Given how fast such content spreads online, its trustworthiness and credibility have become important research areas. We investigate how different factors of social media posts influence their perceived credibility. We generated health-themed short-form social media posts, varied specific aspects of those posts, and deployed the variations on three different online crowdsourcing platforms for credibility assessment. Our quantitative data analysis reveals, for instance, how author professions related to healthcare and science increase the perceived credibility of health-themed posts. Moreover, a higher number of likes and shares increased the credibility in two out of the three platforms. Our qualitative results based on questionnaires highlight personal filtering strategies and critical thinking skills as factors that influence post credibility online. Consequently, our results encourage experts to provide information on social media and to be part of correcting any misinformation as they have higher credibility. Our work strengthens the previous body of work on the credibility of online content in general and acts as a starting point for further studies on social media post content by demonstrating a systematic, crowdsourced, and scalable approach.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Empirical studies in collaborative and social computing**; **Ubiquitous and mobile computing**;

KEYWORDS

Social media, Crowdsourcing, Credibility, Online content, Health claim

ACM Reference Format:
Junhao Li, Miikka Kuutila, Eetu Huusko, Nimantha Kariyakarawana, Marko Savic, Nazanin Nakhaie Ahooie, Simo Hosio, and Mika Mäntylä. 2023. Assessing Credibility Factors of Short-Form Social Media Posts: A Crowdsourced Online Experiment. In *15th Biannual Conference of the Italian SIGCHI Chapter (CHItaly 2023), September 20–22, 2023, Torino, Italy*. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3605390.3605406

1 INTRODUCTION

As the Internet has given society access to information worldwide, people are increasingly making decisions based on online sources such as social media. Social media platforms are proliferating rapidly, and while the benefits of rapid information delivery are undeniable, this development comes with a host of problems as well. The user-generated content fuelling social media contains both truths and rumours, and everything in between, and the credibility of such content can be questioned. For instance, in the American presidential election in 2016, rumours on social media platforms quickly became a national scandal. Further, it has also been shown that, especially during crisis events, people tend to believe and share misinformation using social media [15].

Credibility is defined as the ability to believe in or trust someone [48]. Credibility is one of the oldest communication principles and has recently been a topic of much attention, especially in the context of social media and related research. Both communication scholars and professional communicators have investigated why some communications are perceived as more trustworthy than others [48]. For instance, Xiao et al. have shown that there are multiple factors that affect the credibility assessment from a reader's perspective, such as argument quality or likeability [60]. And trust and credibility are used interchangeably by scholars in their meta-analyses of the credibility of web-based health information [47]. Although credibility assessment is a learnable skill, at least to a

CHItaly 2023, September 20–22, 2023, Torino, Italy

ACM ISBN 979-8-4007-0806-0/23/09.
https://doi.org/10.1145/3605390.3605406

certain degree [21], credibility assessment remains an important societal as well as academic challenge.

In this study, our research question is how different determinants of short-form social media posts affect their perceived credibility. By short-form social media posts, we refer to the newly-popular format of content where the consumer typically sees only a short text piece, along with the author's profile picture (and perhaps a name and a bio – depending on the technical implementation of the platform). We also note that our study is platform-agnostic in the sense, that we do not study a given platform but include these typical elements together in a mock-up of such a social media post. This enables a scalable exploration of the credibility-affecting factors using crowdsourcing platforms.

To address this, we created a rich set of distinct social media posts by combining the use of faces from the Chicago Face Database (CFD) [33] (as the profile pictures of the generated posts), and health claims for nutrition presented by Hans CM and Van der Lans [55]. We additionally augmented the posts with a different number of likes and shares, and the post author's profession, gender, and ethnicity. We deployed a set of posts to three different crowdsourcing platforms, Amazon's Mechanical Turk, Toloka, and Prolific, to obtain a strong sample of post-credibility evaluators.

We make three key contributions:

(1) We contribute to the simultaneous examination of multiple credibility factors of short-form social media posts. This is in contrast to most related work that focuses on the effect of one factor at a time and often on the different types of online content. To this end, we used a computational approach to investigate their relative effects on perceived post credibility;
(2) Our investigation examines credibility factors using samples from multiple online crowdsourcing marketplaces to increase diversity. Although the sample size is limited in certain platforms, it enabled us to provide topical discussion on the suitability of the platforms for this purpose;
(3) We make a novel contribution on how to examine credibility factors through a computational, systematic approach that generalises outside the scope of this article.

As an additional technical contribution, we provide a task batching script for *Prolific* that allows for distributing arbitrary numbers of human intelligence tasks to the crowd workers on the platform in a balanced way. Among others, our quantitative analysis reveals that the profession of the author and the content of the post are particularly relevant to the post's credibility. Additionally, our quantitative analysis also suggests that there are clear differences among the three employed crowdsourcing platforms. Our qualitative analysis identifies *misleading information*, *different types of filtering*, *critical thinking*, *source credibility*, and *consequences of using social media* as particularly intriguing topics that affect people's perceptions about social media credibility. Our study provides a timely investigation of an important topic with a scalable method, providing implications for similar studies in the future.

2 RELATED WORK

2.1 Online Content Credibility

Online content credibility is a pivotal issue in today's society. As people increasingly turn to online sources for their information needs, the content they consume affects their decision-making in important personal as well as societal issues. Information credibility is at the heart of all decision-making [2]. To this end, Allcott and Gentzkow [1] discussed the level of exposure to news about the US presidential election in 2016, discussing how the exposure to fake news has implications for people's beliefs in key governmental functions.

A lot of work has focused on the content itself. Maier et al. [36] have performed experiments on how mass suffering news makes an emotional impact on readers, based on four journalistic elements: story personification, statistical focus, mobilizing information, and photographic depiction of people in need. Their result indicates that journalism using narrative techniques can elicit a stronger emotional response in readers. Similarly, König and Jucks [27] investigated how normal language style vs. aggressive language style, and the professional affiliation (scientist vs lobbyist) of a person arguing in a scientific debate, influence their trustworthiness and the information credibility. They identified that aggressive language leads to lower credibility and normal language to higher credibility. In addition, they also found that a lobbyist was perceived as less trustworthy when compared to a scientist in scientific debates, and yet the credibility of a lobbyist's information was not affected when they delivered scientifically sound and strong arguments. König et al. [29] investigated the effects of enthusiastic language style on the credibility of health information. Their results suggest that an enthusiastic language style leads to lower credibility for scientists and does not have significant effects on lobbyists. Additionally, Lid Rosenholm [31] investigated whether humorous jargon affects users' trust towards health communication in TikTok. Their results from eight interviews with Swedish users aged 19-25 reveal that credibility is heightened when humour is used in health communication on the platform.

Sauls [46] identified that spelling errors play a role in how credibility is perceived. Posts without spelling errors were perceived as significantly more credible than posts with spelling errors. Further, in the case of online retailers, it has been shown that content trustworthiness, expertise, and attractiveness have a positive effect on purchase intention [14]. Kang [25] analysed the credibility of blogs and identified that accuracy and topic (focus) are the key elements to validating the credibility of blog content. Petty et al. [44] conducted a study about personal involvement as a determinant of argument-based persuasion. Their results suggested that an increase in involvement is associated with an increase in the importance of message arguments because people are motivated to hold "correct" and defensible opinions, and they have a better framework for things relevant to the self.

Meanwhile, some other studies investigated factors beyond the content. Many researchers found that the reader demographic factors affect the perceived credibility. Luo et al. [32] introduced the truth-default theory to the context of news credibility. Their results revealed that people often judged news headlines as fake, suggesting a deception bias for news in social media. Metzger and Flanagin [40] conducted a study about the psychological approaches to assess the credibility of online information. In their work, they identified that human information processing activities, demographics, and personality characteristics influence information evaluation, which consequently depends from person to person. Zhou et al. [61] found that the readers' age matters. They gathered 59 older adults aged between 58 and 83 years to examine how eye-catching headlines and emotional images impact their credibility judgements and spreading of health misinformation. Their results indicate that most of their participants would rather trust the misinformation to avoid health risks than doubt it.

Besides that, Maier et al. [35] investigated gender and generational differences in reader reaction to news reports of mass violence in Africa based on the four journalistic elements mentioned before, in which they found women and readers over 30 years old to experience a stronger emotional and charitable response. However, the age gap in emotional response disappeared when presented with either the just-the-facts story or the statistical story. Johnson and Kaye [23] investigated the level of interactivity with 15 sources of political information, as well as the degree of dependence on each source, credibility perceptions, and the strength of interactivity versus the strength of reliance on credibility judgements. They identified how gender, age, education, and income level affect the analysis of credibility. Moreover, they also identified that consumers' trust and belief in the medium affect the credibility of the news and their influence on world view. Fogg et al. [16] also agreed that the readers' education affects credibility perceptions. Johnson and Kaye [22] found that older people, males, and those with a high socioeconomic status tend to be the most critical of the content in general.

Author demographic factors also matter, like the source of the content, which is proven in our other study about social media profile credibility [30]. A study conducted by Johnson and Kaye [22] has found that online publications are more credible than their paper-based counterparts, but participants did not judge each media as credible and suggested that the trust in media is declining. The source's credibility is strongly related to the degree to which people rely on it. Mehrabi et al. [39] conducted a study to determine the factors that influence the credibility perception of the Internet and television. They identified a positive relationship between perceived credibility and the amount of time spent on television viewing and the Internet. In addition, they found television to be rated as more credible than the Internet. Wölker and Powell [59] identified the credibility perceptions of human, automated, and combined news content. In particular, they identified and demonstrated that both automated and combined journal articles are credible alternatives to human-created articles. Moreover, they identified that if the source is considered credible by the reader, then the messages (news) published by the source are considered credible by the readers. Further, Sterrett et al. [50] showed that the reputation of the person sharing a news article positively affects the perceived credibility of the story. Author's gender and ethnicity are also found to affect the perceived credibility [3, 49]

2.2 Social Media Credibility

Social media plays a critical role in today's information environment online. A clear majority of content in social media is user-generated, making credibility an increasingly difficult issue. Also, the reasons why people post content online differ significantly from those of e.g. traditional news media, further exacerbating potential issues with credibility. A key reason to use social media is to reinforce one's identity and to provide a kind of performance online [8]. Furthermore, self-disclosure behaviour, typical on social media platforms, is proven helpful for users to demonstrate their social capital and reinforce their concept of self-authenticity [11–13, 32, 58]. Additionally, research has also focused on how users tend to create their social norms and cluster around groups that support their views, regardless of if the views are factual or not [42, 54]. Overall, promoting critical thinking is generally seen as a positive development to help people understand all kinds of content regardless of whether it is user-generated or not[21].

Other complicating factors exist. Unlike traditional media, social media has built-in popularity indicators that skew readers' opinions, such as comments, retweets and shares. Previous studies put effort into investigating whether this unique component of social media would affect perceived credibility. It is also confirmed that the bandwagon effect, a cognitive bias of people conforming to the crowd and the tendency to follow the majority view, exists in people's behaviours and attitudes toward online information [18, 26]. Similarly, Luo et al. [32] confirmed the effects of endorsement cues in social media (e.g., Facebook likes) on message credibility and detection accuracy. Furthermore, Wijenayake et al. [57] conducted experiments on how a combination of critical and supportive comments on a Facebook news article could influence subsequent readers' perception of the article's trustworthiness as well as their response to it. Their results indicated that people were more inclined to conform to the majority and that conformity is more heightened under critical majorities than under supportive majorities. Results also suggested that initial confidence displayed a significant negative effect on people's conformity behaviour.

Given the prominent role of social media in today's online environment, our study focuses on the credibility factors of a highly popular form of related content: short-form posts, similar to which are popular on phenomenally influential platforms such as Twitter or Instagram.

2.3 Summary of Findings

Based on the literature review, we identify four particularly relevant factors:

- **Reader Demographics:** age [61], gender, education, income level [22, 23, 35, 56], personality traits[40].
- **Author Demographics:** source or Site (publisher) [22, 39, 59], author reputation [14, 50], author's gender and ethnicity [3, 49].
- **Star Rating/Retweets/Likes:** number of positive reviews or comments[18, 26, 57], number of ratings/likes/shares[14, 32], star ratings.
- **Content:** narrative techniques [36], language style [27, 29], humorous jargon [31], information quality, accuracy and relevance [14, 25], spelling errors [46].

In this article, we investigate in detail how **Author Demographics** (profession, age, gender, ethnicity); **Star Rating/Retweets/Likes** (number of likes and shares); **Content** (different evidence) affect credibility. While more exist, we scope our investigation to those above due to their prevalence in most popular social media platforms today that rely on short-form posts.

3 RESEARCH METHOD

3.1 Credibility Rating Task

In the present study, we deployed an online credibility rating task consisting of three health claims on yoghurt. All claims were made with scientific arguments and originated from a paper by Van Trijp and Van der Lans [55]. We combined the claims with different combinations of appearance and profession of authors as well as the number of shares and likes. In our experimental settings, we considered three ethnicities, two genders, and four professions for author profile features. Additionally, we manipulated the posts to have a different number of likes and shares to investigate the impact. As a result, there are 648 ($= 3 * 3 * 3 * 4 * 2 * 3$) unique scenarios in total coming from all combinations of those listed below:

- **Posts claim (3):** Three different health claims about yoghurt from Van Trijp and Van der Lans [55]:
  - **Health claim 1:** Yoghurt helps you build resistance to common diseases, because it contains probiotics.
  - **Health claim 2:** Yoghurt keeps you active and going for longer, because it contains slow-release carbohydrates.
  - **Health claim 3:** Yoghurt helps reduce food intake without making you feel hungry, because it contains added fibres.
- **Likes (3):** Number of likes of the posts. Values: 0, 30, 3,600.
- **Shares (3):** Number of shares of the posts. Values: 0, 57, 5,700.
- **Author profession (4):** The profession of the author of the posts. Values: cook, nurse, professor, social media health influencer.
- **Author gender (2):** The gender of the author of the posts demonstrated by the profile picture. Values: female, male.
- **Author ethnicity (3):** The ethnicity of the author of the posts demonstrated by the profile picture. Values: Asian, Black or White.

For profile pictures, we utilised pictures from the Chicago Face Database (CFD) [33]. CFD is designed to be used in scientific studies and includes high-resolution, standardised pictures of human faces of various ethnicities between the ages of 17 and 65. For each particular model, extensive norming data is supplied. Physical characteristics (e.g., facial size), as well as subjective judgements by independent assessors, are included in this data (e.g., attractiveness). Pictures of 24 individuals were selected based on their different ethnicity and gender. Additionally, we considered other appearance features such as attractiveness, facial expression, etc., which might also influence the result. Therefore, we selected the most average faces in terms of attractiveness by US-based raters provided in the CFD dataset. Finally, we selected all pictures of individuals between the ages of 35 and 45, which was done to ensure that all faces could match all professions, e.g., having a person titled professor and a picture of a 20-year-old creates an obvious mismatch between picture and profession. The selected faces with their professions are shown in Figure 1.

Each scenario was evaluated five times on all the platforms resulting in a total of 3,240 (= $5 * 648$) evaluations. The evaluations were performed by differing amounts of respondents, as shown in Table 2. The number of respondents for the Toloka platform was considerably higher than the rest, as a respondent could stop doing the task at any point, and the rest of their tasks would be given to new participants. The participants were given a post of the health claim with information on the number of shares and likes and the profession and profile picture of the author. After reading the post, the participants submitted their credibility evaluation results using a slider input element (1 to 7). See Figure 2 for an example of how a credibility rating task appeared to the respondents on different platforms.

3.2 Crowdsourcing platforms

According to the results of some previous studies [45, 56], using the arithmetic mean as an aggregation function provides a high level of agreement with the expert label. In our work, we deployed the credibility assessment task to three different paid crowdsourcing and human subject pools online: Amazon's Mechanical Turk (MTurk), Prolific, and Toloka. Screenshots of the task on different platforms can be seen in Figure 2.

MTurk is a commonly used crowdsourcing platform, used in industry as well as in academia [43]. Typical tasks in MTurk include short tasks, such as data labelling, sentiment analysis, and information searches online.

Prolific is a newer platform, founded in 2014, that focuses on behavioural research. As such, it is more akin to a human subject pool than a typical crowdsourcing platform.

Finally, we used Toloka, which offers globally millions of workers mainly for industrial data labelling and collection purposes [9].

For Toloka participants, we paid 4 cents for each social media credibility rating, and 40 cents for completing the follow-up survey. Additionally, we paid one and two-dollar bonuses for long and rational answers to the open-ended questions in the survey. For MTurk, we paid 4 cents for the ratings and $1.50 for the follow-up survey, while on Prolific, we paid $0.33 per batch of nine ratings and $1.67 for the follow-up surveys.

Typically, studies elicit data from only one of these platforms. We argue that choosing to obtain response data from all three platforms is a particular strength in our study, as it helps increase diversity among our participants and increases the external validity of our results.

Prolific offers a human subject recruiting platform where participants must be sent to an external URL to complete tasks. However, in a research case like ours, where hundreds of permutations of a post are available, a batching solution to provide the permutations to any arbitrary number of workers is needed. MTurk and Toloka offered such a solution, but for Prolific, we needed to implement one.

To this end, in this article, we also contribute an open-source software for batching tasks with Prolific (or indeed any human subject pool that allows sending participants to an URL), available at GitHub: (URL_REDACTED_FOR_ANONYMISATION). The tool ensures each subject gets an equal amount of tasks, and that the tasks are diverse with minimised repetition. The former aspect is important as it gives each participant an equal say, while the latter is significant as random sorting can result in some workers seeing repetitive tasks, which can affect their engagement level and response quality.

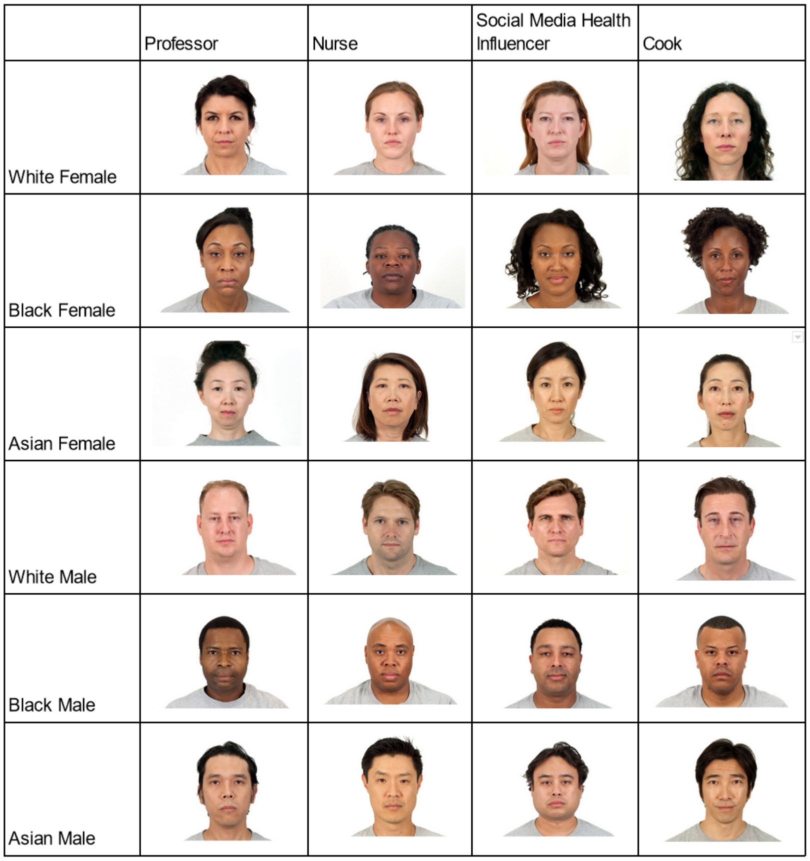


Figure 1: Faces and professions used in our study.

To illustrate the batching mechanism, let us consider a stack of N tasks that need to be sorted into M batches. Each worker will be entrusted with completing one batch of tasks B_i, with $i \in \{1, ..., M\}$. To ensure all batches are evenly sized, an upper limit $UL = ceil(N/M)$ is established, which indicates the maximum task capacity of each batch. The batches are initially empty and are filled up following this loop:

(1) Task t is popped from the task's stack.
(2) f_{sim} is a similarity function that compares each element of two tasks and returns a score that is incremented for every identical element. A cumulative similarity score is computed between t and each B_i.

$$S_{t,i} = \sum_{b \in B_i} f_{sim}(t, b) \qquad \forall i \in \{1, ..., M\}$$

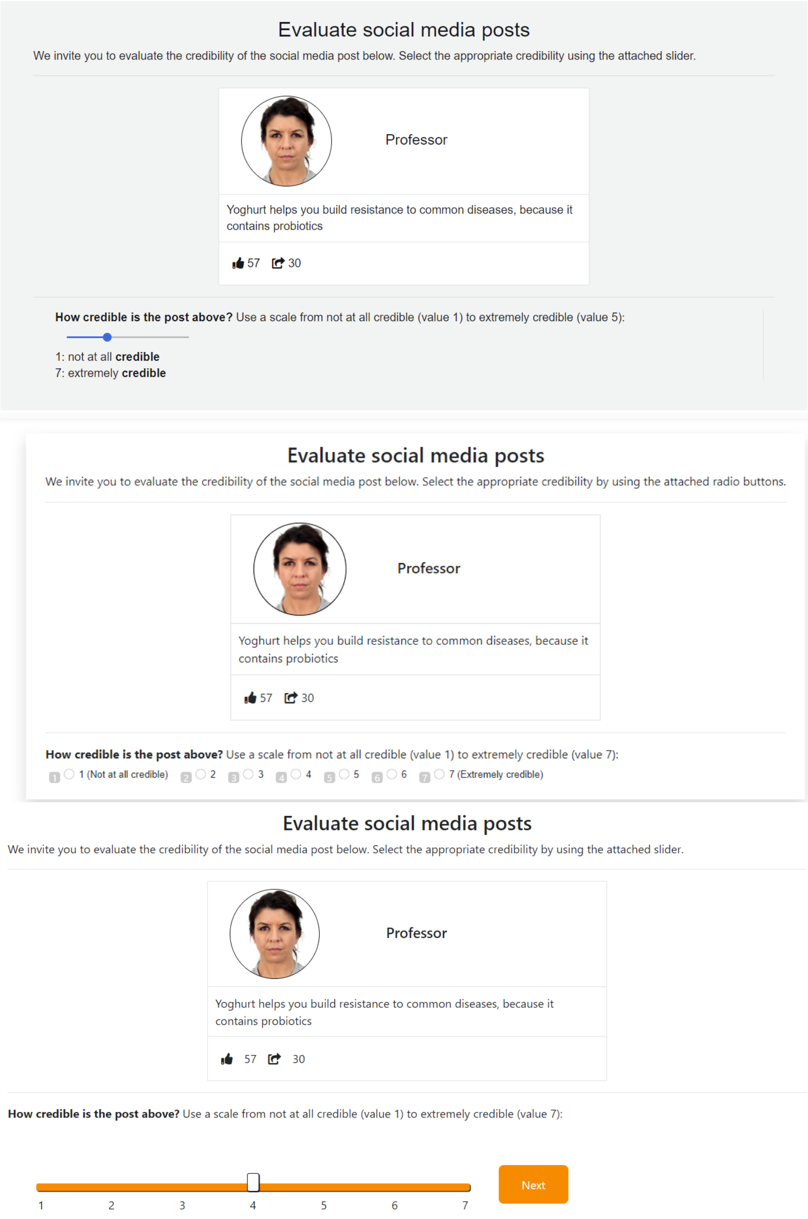
(3) Ignoring those batches which are larger than UL, t is sorted in the batch with the lowest cumulative similarity score, $t \rightarrow min_{i \in \{1,...M\}}(S_{t,i})$
(4) The process is repeated until all tasks are sorted into batches.

Each participant needs to rate at least one post, and they would not rate posts with identical elements rated before. Toloka and MTurk provide this function. For Prolific, we used the above-mentioned open-source software to achieve that. The task procedure is as follows:

(1) Accept the task
(2) Redirect to the credibility rating task page as shown in Figure 2
(3) Read the post and give a credibility rating.
(4) Click next if the participant wants to rate another post. The next button is not shown in MTurk and Toloka because the platform provides it.

3.3 Follow-up Survey

After completing the credibility rating task, described in Section 3.1, we invited the respondents to answer a follow-up survey consisting of collecting background and demographic information, as well as opinions related to social media usage and credibility. The demographic information includes age, gender, nationality, ethnicity, marital status, education level, employment status, and annual income level of the respondents. As the foundation of our survey, we used a demographics questionnaire by Pew Research [5]. As there is no agreed-upon international standard for ethnicity classification [7], we used the suggested classification from the National Content Test 2015 [24, 37]. The demographic information is gathered to investigate the diversity among the Crowdsourcing platforms



Evaluate social media posts

We invite you to evaluate the credibility of the social media post below. Select the appropriate credibility using the attached slider.

Professor

Yoghurt helps you build resistance to common diseases, because it contains probiotics

57 30

How credible is the post above? Use a scale from not at all credible (value 1) to extremely credible (value 5):

1: not at all **credible**
7: extremely **credible**

Evaluate social media posts

We invite you to evaluate the credibility of the social media post below. Select the appropriate credibility by using the attached radio buttons.

Professor

Yoghurt helps you build resistance to common diseases, because it contains probiotics

57 30

How credible is the post above? Use a scale from not at all credible (value 1) to extremely credible (value 7):

1 ○ 1 (Not at all credible) 2 ○ 2 3 ○ 3 4 ○ 4 5 ○ 5 6 ○ 6 7 ○ 7 (Extremely credible)

Evaluate social media posts

We invite you to evaluate the credibility of the social media post below. Select the appropriate credibility by using the attached slider.

Professor

Yoghurt helps you build resistance to common diseases, because it contains probiotics

57 30

How credible is the post above? Use a scale from not at all credible (value 1) to extremely credible (value 7):

Figure 2: Credibility rating task on all three platforms. From top-down: MTurk, Toloka, and Prolific.

we use and might be in use for future research. We also asked the respondents about what social media platforms they use.

We asked the respondents for their opinion on 12 items affecting the credibility of social media posts, which are shown in Table 1 with a seven-point ordinal scale. We selected these 12 items based on our findings from the literature review, summarised in the related work section. Additionally, we asked the respondent to rate these questions on a seven-point scale: ”How do you rate yourself in terms of assessing the credibility of social media posts, i.e., your critical social media reading skills?” and ”In your own opinion, how credible do you think social media, in general, is now as an information source?". We also had two open-ended questions, one requesting general opinions on the matter of social media credibility, including those affecting the respondent or society, and the second open-ended question asking respondents to elaborate on how credible social media posts are in general.

3.4 Statistical Analysis

We used the R package *Ordinal* to perform a cumulative link mixed model (CLMM), an extension of the cumulative link model (CLM), analysis of the relationship between the aforementioned features

Table 1: Items are rated on a seven-point ordinal scale affecting the credibility of a social media post. The question is formulated as "In your opinion, how much do the following items affect the perceived credibility of a single social media post?".

Author's profession
Author's gender
Author's age
Availability of multiple sources to confirm
Its spelling and grammar
Tone of writing
Appearance of the post
Trustworthiness towards author
The social media outlet
Income level of the reader
Education level of the reader
Amount of time spent on social media

and participant-evaluated credibility results. The CLMM model allows us to include random effects to model group heterogeneities. As a random effect, we specified participant (WorkerID) to account for individual differences in our model. The CLMM model has an ordinal outcome (response) variable Y_i, which is the numerical human-evaluated credibility of posts in this study. All demographic factors above were categorical predictors. The syntax of the final CLMM model in R is as follows: $Credibility \sim Profession + Shares + Likes + HealthClaim + Gender + Ethnicity + \frac{1}{WorkerID}$.

3.5 Qualitative Analysis

We coded the open-ended follow-up survey items following recommendations of Braun and Clarke [4] for thematic analysis, with the following modifications. The first two authors both open-coded 50% of the responses separately. The codes were then compared and discussed with four authors. Now, the codes were refined and merged together into one set. After this, both of the first two authors again applied defined codes to all responses individually. The authors met weekly to share the code updates and merge the differences through a discussion. The process was continued until no conflicts with codes remained, as also suggested in McDonald et al. [38].

4 DATA ANALYSIS

4.1 Participants

We invited all the participants to do a follow-up survey to gather additional information. In the end, we received 16 responses from the original 36 participants from MTurk, 161 responses from 418 Toloka participants, and 293 responses from the 360 Prolific participants. The average age of the participants was 35.7 (24-69) in MTurk, 29.7 (18 - 70) in Toloka, and 43.8 (21 - 77) in Prolific.

Of the MTurk users, 11 reported themselves as male and 5 as female; of the Toloka users, 103 reported themselves as male, 57 as female, and one as "LGBT"; and of the platform Prolific users, 118 respondents reported themselves as male, 174 as female, and one as non-binary.

In MTurk, three participants were from India and the rest were from the United States. In Toloka, 167 participants were from Russia, 33 from Turkey, and 21 from both Brazil and India, while the rest of the 176 participants were from other countries. In Prolific, 233 participants were from the United Kingdom, 23 from the United States, 8 from Italy, and the rest were from other countries.

In MTurk, one participant reported themselves as holding a master's degree, 11 reported themselves as holding a bachelor's degree, and 4 had a high school diploma. In Toloka, 28 participants reported themselves as holding a master's degree, 69 reported having or being in the process of getting a bachelor's degree, 40 as having a High school diploma, 18 as having a Professional degree, with the rest (6) reported themselves as not having any qualifications. For Prolific, 42 participants reported themselves as holding a master's degree, 123 a bachelor's degree, 92 a high diploma, 13 a professional degree, and 10 a doctoral degree.

Table 2 shows the information on the respondents. With regards to nationalities, MTurk and Prolific focus more on western countries, whereas respondents from the Toloka platform are more widespread throughout the world. Prolific users were the most likely to complete the follow-up survey, with over eighty percent responding, whereas respondents from the MTurk and Toloka platforms responded to it around 40 percent of the time. With respect to gender, MTurk and Toloka had around two male respondents to every female respondent, but for Prolific, this was almost reversed with females being the clear majority. Bachelor's degrees were the most commonly reported educational achievement on all platforms. We also gathered responses about the participants' self-scoring critical reading skills and trust in social media. It was seen that the participants tended to give themselves higher scores on their own critical reading skills on all three platforms. The average scores are all higher than 5. On the other hand, the participants all gave lower scores for their trust in social media. The Toloka participants trust social media the most, with the highest average score of 4.47, while the average score of MTurk participants is 3.69 and Prolific is 3.24.

4.2 Quantitative Credibility Analysis

The CLMM was used to examine relationships between post factors and the perceived credibility of posts. Three models with data from different platforms are shown in Tables (3, 4 and 5). In the tables, the estimate is the estimated relationship between the fixed effect term on the same row and rated credibility. The p-value describes how likely it is that the relationship (estimate) is by random chance, with p-values of less than 0.05 interpreted as a statistically significant relationship. The standard error is the average distance of observations from the regression line. Lastly, the Z value is the estimate divided by the standard error. Cohen's d value is a statistical measure to quantify the effect size.

All our variables are dichotomous, which means we model them with dummy variables [20]. Each variable represents a dichotomous fact. For example, if the evaluated social media post was made by a professor, then our variable Profession_Professor is set to 1, and all other Professions _* dummies are set to zero. Using dummy variables requires always setting one dichotomous fact as the default case that has no statistical values, as all other dummies are compared

Table 2: Information on the participants on all platforms. The percentage of the total for a specific platform is given in brackets.

	MTurk	Toloka	Prolific
Main task respondents	36	418	360
Follow up survey respondents	16 (44%)	161 (39%)	293 (81%)
Average age	35.7	29.7	43.8
Minimum age	24	18	21
Maximum age	69	70	77
Gender:			
Male	11 (69%)	103 (64%)	118 (40%)
Female	5 (31%)	57 (35%)	174 (59%)
Other	0	1 (1%)	1(<1%)
Ethnicity:			
Asian	3 (19%)	42 (26%)	17 (6%)
Black	1 (6%)	15 (9%)	9 (3%)
White	12 (75%)	70 (43%)	261 (89%)
Other	0	34 (21%)	6 (2%)
Top 3 nationalities:	United States - 33 (92%)	Russia - 167 (40%)	United Kingdom - 233 (65%)
	India - 3 (8%)	Turkey - 33 (8%)	United States - 23 (6%)
		India & Brazil - 21 (5%)	Italy - 8 (2%)
Rest of nationalities:	0 (0%)	176 (42%)	96 (27%)
Education:			
Doctorate	0	0	10 (3%)
Master’s degree	1	28 (17%)	42 (14%)
Bachelor’s degree	11 (69%)	69 (43%)	123 (42%)
High school diploma	4 (25%)	40 (25%)	92 (31%)
Professional degree	0	18 (11%)	13 (4%)
Other qualifications or none	0	6 (4%)	13 (4%)
Self-scoring average results:			
Critical reading skill	5.69	5.51	5.31
Trust in social media	3.69	4.47	3.24

against the default case. These defaults are also shown in Tables (3, 4 and 5)

We find that professions and post credibility have a statistically significant relationship in our models, as shown in Tables 3, 4 and 5. In all datasets, Nurses and Professors are the most credible as they have the highest positive estimates on post credibility. For both Amazon and Prolific platforms, all professions have a p-value of less than 0.001 and a Cohen’s d absolute value above 0.5. While for Toloka, one is significant at each of the 0.05, 0.01, and 0.001 levels. But according to Cohen’s d values of Toloka, only the professor reached the small effect size. The perceived credibility of different professions from highest to lowest using the different platforms was as follows:

- MTurk: Nurse, Professor, Cook, and Social media health influencer
- Toloka: Professor, Nurse, Social media influencer, and Cook
- Prolific: Professor, Nurse, Cook, and Social media influencer

The increase in the number of shares and likes should, according to related work, increase the estimate of the post with credibility. Yet, this varies between platforms, providing interesting insights into the use of such platforms themselves. In MTurk and Toloka, the behaviour is as expected, as shown in Tables 3 and 4. But Cohen’s d values suggested that the highest number of likes only reached the medium effect size in MTurk and the small effect size in Toloka. However, in Prolific, we find no statistically significant effect from the number of likes and shares of the post, see Table 5.

In our dataset, most gender and ethnicity-related factors were not statistically significant predictors of credibility. The only statistically significant predictor related to ethnicity is black ethnicity for the MTurk platform shown in Table 3, where the estimate of credibility was negative when compared to the default Asian ethnicity, and Cohen’s d value suggested the effect size did not reach medium size. Regarding gender, the only statistically significant predictor is in the Prolific platform, shown in Table 5, where the estimate is those female post authors were seen as more credible than males, and Cohen’s d value suggested the effect size did not reach small size.

Moreover, the claim made in the posts affects the rater’s judgement of credibility, as all claim-related variables have statistically significant relationships with credibility. Health claim 1, formulated as *"Yoghurt helps you build resistance to common diseases, because it contains probiotics"*, is the most credible one for all platforms, as estimates for claims 2 and 3 have negative estimates. Health claim 3, *"Yoghurt helps reduce food intake without making you feel hungry, because it contains added fibres."* is the least credible claim with the most negative estimate for all platforms. However, Cohen’s d values suggested different effect sizes in different platforms. In MTurk, both claim 2 and claim 3 did not reach a small effect size, and in Toloka, claim 3 reached a small effect size. While in Prolific, claim

2 reached a small effect size, and claim 3 reached a medium effect size.

Table 3: CLMM Result of Amazon Mechanical Turk. P-values are marked with * <0.05, ** <0.01, and * <0.001.**

Credibility factor	Estimate	std. error	z value	*p*
Default: Profession_Cook				
Profession_Nurse	0.98750	0.08888	11.111	<2e-16 ***
Profession_Professor	0.97107	0.09053	10.726	<2e-16 ***
Profession_Influencer	-0.49618	0.09068	-5.472	4.46e-08 ***
Default: Shares_0				
Shares_30	0.13993	0.07740	1.808	0.07063
Shares_3600	0.67329	0.07853	8.574	<2e-16 ***
Default: Likes_0				
Likes_57	0.30720	0.07710	3.984	6.77e-05 ***
Likes_5700	0.93150	0.07956	11.709	<2e-16 ***
Default: HealthClaim1				
HealthClaim2	-0.22445	0.07799	-2.878	0.00401 **
HealthClaim3	-0.30914	0.07778	-3.975	7.05e-05 ***
Default: Gender_Female				
Gender_Male	-0.07223	0.06318	-1.143	0.25291
Default: Ethnicity_Asian				
Ethnicity_Black	-0.60996	0.07843	-7.777	7.43e-15 ***
Ethnicity_White	-0.01951	0.07685	-0.254	0.79962

Table 4: CLMM Result of Toloka. P-values are marked with * <0.05, ** <0.01, and * <0.001 .**

Credibility factor	Estimate	std. error	z value	*p*
Default: Profession_Cook				
Profession_Nurse	0.25459	0.09652	2.638	0.008346 **
Profession_Professor	0.40406	0.10035	4.026	5.67e-05 ***
Profession_Influencer	0.22970	0.09788	2.347	0.018941 *
Default: Shares_0				
Shares_30	0.28173	0.08516	3.308	0.000938 ***
Shares_3600	0.33242	0.08457	3.931	8.47e-05 ***
Default: Likes_0				
Likes_57	0.26385	0.08510	3.100	0.001932 **
Likes_5700	0.62223	0.08576	7.256	4.00e-13 ***
Default: HealthClaim1				
HealthClaim2	-0.29960	0.08575	-3.494	0.000476 ***
HealthClaim3	-0.44183	0.08604	-5.135	2.82e-07 ***
Default: Gender_Female				
Gender_Male	-0.06678	0.06964	-0.959	0.337564
Default: Ethnicity_Asian				
Ethnicity_Black	-0.02106	0.08501	-0.248	0.804318
Ethnicity_White	-0.04025	0.08432	-0.477	0.633118

4.3 Quantitative Survey Analysis

Figure 3 demonstrates the distribution of scores for the 12 items participants were asked about. The importance of these items affecting the credibility of social media posts is ranked in the following order: Availability of multiple sources to confirm; Trustworthiness towards the author; Its spelling and grammar; Author's profession; Tone of writing; Education level of the reader; Appearance of the post; The social media outlet; Amount of time spent on social media; Author's age; Income level of the reader; and the Author's gender. Similar to our CLMM results, the author's profession is considered an important factor for the credibility of social media posts, while the author's gender is not. There were statistically significant differences between the items, as confirmed with a Kruskal Wallis test (p <0.01). We refrain from including a full pairwise comparison table, as Figure 3 depicts a clear visual overview of the differences between the individual factors.

Table 5: CLMM Result of Prolific. P-values are marked with * <0.05, ** <0.01, and * <0.001 .**

Credibility factor	Estimate	std. error	z value	*p*
Default: Profession_Cook				
Profession_Nurse	0.95775	0.09193	10.418	<2e-16 ***
Profession_Professor	1.06448	0.09338	11.400	<2e-16 ***
Profession_Influencer	-1.02791	0.09425	-10.906	<2e-16 ***
Default: Shares_0				
Shares_30	-0.08282	0.07964	-1.040	0.29839
Shares_3600	-0.04904	0.07945	-0.617	0.53708
Default: Likes_0				
Likes_57	0.09535	0.07942	1.201	0.22992
Likes_5700	0.11295	0.07950	1.421	0.15538
Default: HealthClaim1				
HealthClaim2	-0.58271	0.08099	-7.195	6.24e-13 ***
HealthClaim3	-0.89931	0.08130	-11.062	<2e-16 ***
Default: Gender_Female				
Gender_Male	-0.17697	0.06818	-2.596	0.00944 **
Default: Ethnicity_Asian				
Ethnicity_Black	0.03268	0.08286	0.394	0.69326
Ethnicity_White	0.11038	0.08294	1.331	0.18327

4.4 Qualitative Survey Analysis

In the follow-up survey, we probed people's opinions on the issue of social media credibility at a higher level. The following themes describe the participants' sentiments:

Misleading information. Fake news, false information, misinformation, misleading information, and similar words are mentioned frequently in participants' responses. Specifically, participants expressed concern about the prevalence and propagation of such content on social media platforms. According to some participants, the amount of misleading information on social media is significantly higher than in traditional media like newspapers. Furthermore, participants believe there are people or organisations intentionally creating and disseminating misleading information with the aim of attracting attention or stirring up hate and racism. This was particularly concerning for the participants, as they recognised the potential harm that such content could cause to individuals and society as a whole.

> *Most of social media is quite fake, especially how people present themselves. Also, there is a lot of fake news going around, so, in general, social media is often not credible*
> – Female, Germany, aged 38

User originated filtering. This theme concludes the participants' responses towards the user-originated filter of the information propagation process from the produced to that perceived as true. Authors' selectively posting strategies, such as the use of

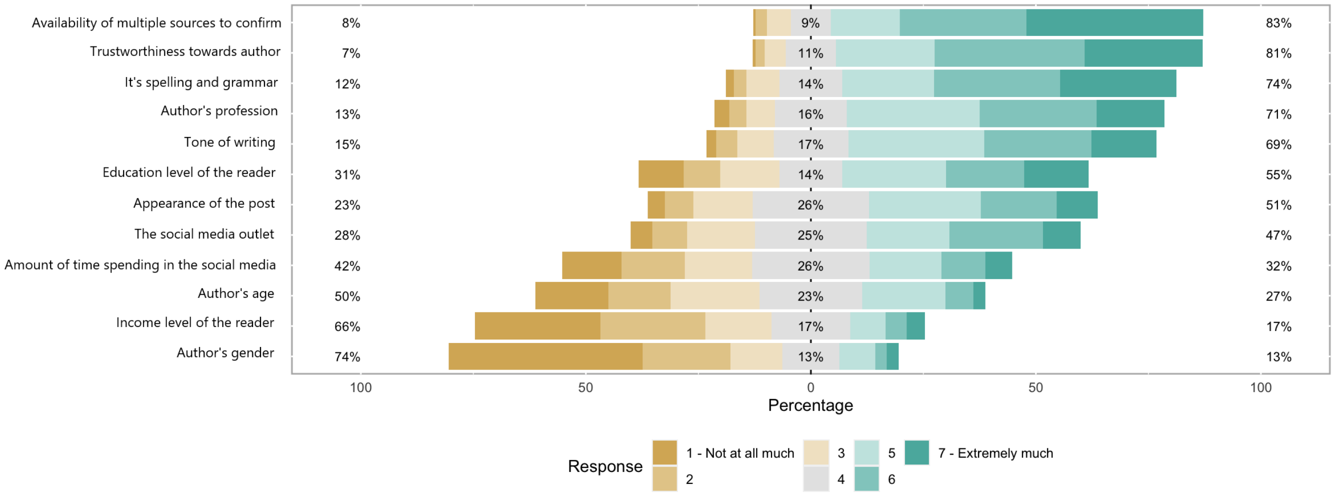


Figure 3: Questionnaire responses to how much different factors were perceived to affect the credibility of a short-form social media post.

click-bait headlines or emotionally charged language, can influence what information users see and engage with. Furthermore, users' subjective willingness to trust information and the position they stand in can affect what information they trust. For example, participants discussed how individuals with certain political or ideological beliefs might be more likely to trust information that aligns with their views, even if it lacks evidence or is misleading. This can create a "filter bubble" effect, where individuals are only exposed to information that reinforces their existing beliefs and biases[10].

> *I don't trust what is posted on social networks because most of them have a bias, an ideology behind them. Besides, there is a lot of selectivity of what is posted.*
> – Female, Brazil, aged 27

Platform filtering. Participants recognised that these algorithms are designed to personalise the user experience and that they use a variety of factors, such as interests, past behaviour, and demographics, to determine what content is shown. However, they were concerned about the potential for social media platforms to prioritise content that benefits their own interests, such as advertisements, over the needs and interests of users. This was seen as a potential conflict of interest, where platforms may promote content that generates revenue or meets other business purposes rather than providing users with accurate and trustworthy information. Furthermore, some people worried platforms and governments could also filter the transmission of certain information due to policies or their own interests. This was seen as a potential threat to freedom of speech and democratic values. Participants recognised that while there are limits to free speech, such as in cases of hate speech or incitement to violence, transparency and accountability in how social media platforms and governments regulate the flow of information matters to help to restore public trust in digital platforms[34].

> *Most of the social media posts are paid, so they are as credible as their customers want them to be. It is easy to push an agenda on social media or manipulate others.*
> – Male, Hungary, aged 31

Critical thinking ability. Many participants expressed confidence in their critical thinking ability and believed that they were able to discern reliable information from misleading content. However, some participants also acknowledged that it could be challenging to differentiate between accurate and inaccurate information, particularly given the vast amount of content available on social media platforms. Unsurprisingly, participants tended to view individuals with lower levels of education as being more vulnerable to misleading information. Most participants believed that critical thinking ability is important when using social media platforms that are flooded with misleading information. They emphasised the need for users to actively evaluate the accuracy and reliability of the information rather than simply accepting everything they see on social media as true.

> *People obviously try to influence people through social media. Youngsters seem to take it as written, they do not check the veracity of any of the opinions put on there, they just jump on the bandwagon and believe what their mates write.*
> – Female, United Kingdom, aged 60

Source credibility. Several respondents highlighted the importance of the source of information in determining its credibility. Participants used various terms such as source credibility, content credibility, author credibility, and source reliability to describe their assessments of post credibility. In particular, participants emphasised the importance of the credibility of the author or publisher in determining the credibility of a social media post. They believed that if the author/publisher is trustworthy, the post is more likely to be credible, regardless of the social media platform on which it is shared. This highlights the role of reputation and trustworthiness in social media, where users rely on the perceived credibility of authors/publishers to assess the reliability of the information.

> *I think social networks can be trusted. They are trustworthy. Not all of course, so it is worth considering the source or organisation.*
> – Male, Russia, aged 34

> *If we look at reliable sources, we can find the right information. We should not rely on information that has no reliable source.*
> – Male, Turkey, aged 23

Consequences of use. Participants expressed their concerns about the potential harm caused by misinformation/disinformation. They believed that using social media without being aware of a post's credibility can have serious implications for individuals and communities. For individuals, it might lead to different consequences, such as personal breakdowns, mental health issues and substance abuse/misuse, which has increased a lot during the COVID-19 pandemic. At the community level, the consequences are even more serious. It can even deliberately incite acts of hatred against specific ethnic groups. It has been proven that anti-Asian hate crimes in America have skyrocketed since COVID-19[19].

> *With so many people able to post freely, I think it's really difficult to monitor and filter out posts that aren't credible. I think this impacts mostly on people who aren't able to assess the posts themselves, mostly people who are more vulnerable in society, so I think it's damaging. I'm not sure what the answer is - freedom of speech, yes, but not at the expense of the health or well-being of others.*
> – Female, United Kingdom, aged 34

5 DISCUSSION

5.1 Credibility Factors in Social Media

In this study, we set out to investigate how different factors affect the perceived credibility of short-form social media posts. Our findings confirm that these factors affect readers' evaluation results to varying degrees.

Based on the quantitative analysis of our results, the profession of the author affects the credibility of a social media post on health claims significantly, in line with findings of König and Jucks [28]. It is also suggested in our qualitative analysis that people tend to trust posts from trustworthy authors. The professions of Nurse and Professor ranked higher than other professions, and more specifically, Professors ranked higher than Nurses in both Toloka and Prolific platforms. In MTurk, nurses were the highest-ranked profession, but the difference in the credibility estimates was negligible (0.99 vs 0.97). This might also be caused by the small number of MTurk participants. Interestingly, the profession of a social media health influencer was the least credible in MTurk and Prolific. However, in Toloka, this profession was ranked higher than that of a Cook. While further research is warranted on the use of these platforms, we can speculate that certain demographic factors could explain this result. For instance, the average age we managed to recruit was lower, and the background was more diverse in Toloka than on the MTurk and Prolific platforms.

Our results contrast the findings of Spence et al. [49], who identified that the ethnicity of the author influences the credibility of a social media post. After making six comparisons related to ethnicity, in only one of those comparisons were statistically significant differences found. With MTurk, Black individuals as post-authors were deemed less credible than Asians. Yet, this finding was not replicated in the Prolific or Toloka platforms, and it might be caused by the limited sample size of MTurk as well.

Previous work has also indicated that male authors were deemed more credible compared to their female counterparts [3]. The results of our study, however, do not fall in line with these findings. In only one platform (Prolific) out of the three, we found a statistically significant benefit between genders, and the difference indicates that women are deemed more credible. This finding did not replicate in the two other platforms. We note that the samples of respondents are different from Armstrong and McAdams [3], where their respondents were from undergraduate classes in a university in the southeastern US, while in our Prolific sample, the majority of respondents are from the UK and female.

The results of our study are in line with previous work by Kang [25], where likes on social media directly increase the estimated credibility of the post for the MTurk and Toloka platforms. The same holds true for the Prolific platform, though the effect was not statistically significant. On social media, the number of shares also increases the credibility of the post on the MTurk and Toloka platforms. However effect could not be found in the Prolific platform ratings, as the estimates for 30 and 3,600 shares have a minus sign for the base case of 0, and they are not statistically significant.

In Tables (3, 4 and 5), health claims are significant predictors of credibility in all the models for all the platforms. Moreover, the rankings are consistent, with claim 1 being the most credible and claim 3 being the least credible on all platforms. This suggests that claim and argument types play a significant role in assessing the credibility of social media posts. Prior work has also noted the importance of the credibility of claims, e.g., [6, 51]. Thus, more experiments should be done to explore how claims and arguments affect the credibility of social media health claims.

5.2 Crowdsourcing Platforms for Credibility Evaluation

Previous research has investigated the pros and cons of crowdsourcing platforms. Oppenlaender et al. [41] conducted an experiment, and their results revealed clear differences between the workers available on two commonly used platforms, MTurk and Prolific. Despite the fact that our sample size is not very large, our investigation that used all these platforms provides some timely insights on their differences, with practical considerations for researchers:

Crowd worker diversity. In our experiment, it was clear that Toloka and Prolific were more diverse compared to MTurk. Prolific participants came from different countries, mostly in Europe, while Toloka participants came from different continents such as Africa, Asia, Europe, the Americas, and Australia. In addition, MTurk participants were limited to citizens of two countries, and most were American. Despite the fact that participants of all three platforms have a large percentage of white people, the diversity of participants was highest in Toloka, followed by Prolific and MTurk. We need to mention that the sample size of MTurk is small compared to Prolific and Toloka. Therefore this might not be exactly true that MTurk is less diverse than other platforms.

Back-end usability. While implementing our experiments on these platforms, we noticed that there are practical usability differences among them. Toloka provided the most convenient and usable back-end implementation experience, with granular controls for task distribution and different ways of uploading the permutations to the platforms. Prolific only supports the recruitment of subjects, for which we had to implement a full third-party task batching solution. Finally, MTurk documentation is still rather poor, and the platform's back-end functionality is rather confusing. While these findings are anecdotal, they provide pointers to researchers wishing to reproduce similar work.

5.3 Public Opinion on Social Media Credibility

Misleading information is a threat is an opinion now not only shared by society at large, but our participants agreed with the sentiment too. They expressed their concerns about the quantity and speed of dissemination of misleading information in social media far exceeding those of traditional media. Moreover, some of them believe there are more individuals or organisations intentionally spreading misleading information on social media in order to profit or stir up public opinion. From cyberbullying to inciting racism mentioned by participants suggested that misleading information is now a largely acknowledged state in social media.

The *ubiquitous information cocoon* is communication where we hear only what we choose and what comforts and pleases us [52]. Related to this concept, some of our participants suggested that social media users would be easier to be trapped in such information cocoons. Summarising their opinions in our qualitative analysis, one reason might be user-originated filtering. Specifically, our analysis of user-originated filtering in our results suggests that users tend to selectively share content that attracts more attention and are more likely to follow users who post content they are interested in. Another reason might be platform filtering. For example, participants were aware that recommendation algorithms generally would only show posts on certain topics they used to enjoy. Moreover, participants also suggested that social media platforms are likely to promote content for their own benefit.

Some participants shared their thoughts on Combating Misleading Information. Our results revealed participants' thoughts of critical thinking skills' importance for social media users, which would decrease the number of victims and slow down the dissemination speed of misinformation. Previous studies have proven that this skill is learnable and can be improved to a certain degree [21]. Several participants also highlighted the reputation of the author when assessing credibility and suggest to train users to be able to verify it. Moreover, participants expressed the potential for content moderation by platforms/governments to stop the spread of misinformation. However, they were also concerned it might be dangerous if such power to manipulate information is in certain people/groups' hands.

5.4 Limitations and Future Work

As respondents were paid for their answers in crowdsourcing platforms, completing them in a fast manner and maximising income is a potential threat to validity [17]. However, we mitigated this by using filters in all used crowdsourcing platforms to recruit only the top-performing workers for our tasks. Further, we note that these platforms are now used commonly across different academic fields and have been shown to produce ecologically valid data [53].

Additionally, how the participants consumed our posts in the simulated environment differs from how people consume social media on the Internet. People typically scroll when browsing social media, and in that case, people would not see the authors' titles on most social platforms even if they have added their titles to the bios. Therefore, our results do not generalise to people scrolling social media but to people reading a post in a way where they have all the details at their disposal (this depends on the platform implementation).

Lastly, one limitation in our work is the health claim used to gauge respondents' credibility ratings. There are three different versions of a health claim without any theoretical background on why one should be preferred over another. Additionally, the health claims involved in this research are limited to yoghurt. Thus, the claims might not be generalisable to all health claims. In future work, we intend to study the effect of argument or claim structure on the credibility of social media posts.

For further additional work, different platforms other than MTurk, Toloka, and Prolific can be used to gather data. The data can be compared and contrasted with the data gathered during this research. Furthermore, this can be expanded to domains other than health claims and compare and contrast whether the results can be replicated or whether people react differently to different domains. With a higher amount of respondents, biases related to ethnicities can also be examined more closely by taking respondents' reported ethnicity into account in the regression models. We also consider having a moderating variable on how much interest people have in healthy food information, as the results may vary. Moreover, we noticed that some participants expressed concerns about the credibility of social media due to the prevalence of misinformation. It could be further explored how different levels of general scepticism or distrust in social media affects people's credibility assessment process.

6 CONCLUSIONS

In this paper, we investigated how different factors affect the perceived credibility of online social media posts using a scalable approach and three different crowdsourcing platforms. Additionally, we contribute a versatile script for task distribution on crowdsourcing marketplaces that do not support task batching natively. Our results indicating the limited impact of gender and ethnicity in credibility evaluations contradict prior work, warranting further research in this area. However, most of our results regarding factors influencing credibility align well with past work: profession, claim, likes and shares affect credibility. Crowdsourcing platforms, on the other hand, are a relatively new participant recruitment channel for studies like this. To this end, we provide a novel case study by using not one but three sources simultaneously. Put together, our results and discussion provide a timely piece to the puzzle of how short-form social media credibility is construed and demonstrate a scalable and systematic approach for credibility investigation.

ACKNOWLEDGMENTS

This research is connected to the GenZ strategic profiling project at the University of Oulu, supported by the Academy of Finland (project number 318930), Biocenter Oulu, the Strategic Research Council (SRC), established within the Academy of Finland (Grants 335625, 335729), and Academy Research Fellow funding by Academy of Finland (Grants 349637 and 353790).

REFERENCES

[1] Hunt Allcott and Matthew Gentzkow. 2017. Social media and fake news in the 2016 election. *Journal of economic perspectives* 31, 2 (2017), 211–36.

[2] Majed Alrubaian, Muhammad Al-Qurishi, Atif Alamri, Mabrook Al-Rakhami, Mohammad Mehedi Hassan, and Giancarlo Fortino. 2018. Credibility in online social networks: A survey. *IEEE Access* 7 (2018), 2828–2855.

[3] Cory L Armstrong and Melinda J McAdams. 2009. Blogs of information: How gender cues and individual motivations influence perceptions of credibility. *Journal of Computer-Mediated Communication* 14, 3 (2009), 435–456.

[4] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[5] Pew Research Center. 2015. Pew Research Center Demographic Questions Web or Mail Mode 12-29-2015. https://assets.pewresearch.org/wp-content/uploads/sites/12/2015/03/Demographic-Questions-Web-and-Mail-English-3-20-2015.pdf, Accessed: 13-8-2021.

[6] Nyamragchaa Chimedtseren, Bridget Kelly, Anne-Therese McMahon, and Heather Yeatman. 2020. Prevalence and Credibility of Nutrition and Health Claims: Policy Implications from a Case Study of Mongolian Food Labels. *International Journal of Environmental Research and Public Health* 17, 20 (2020), 7456.

[7] Roxanne Connelly, Vernon Gayle, and Paul S Lambert. 2016. Ethnicity and ethnic group measures in social survey research. *Methodological Innovations* 9 (2016), 2059799116642885.

[8] Judith Donath and Danah Boyd. 2004. Public displays of connection. *BT technology Journal* 22, 4 (2004), 71–82.

[9] Alexey Drutsa, Valentina Fedorova, Dmitry Ustalov, Olga Megorskaya, Evfrosiniya Zerminova, and Daria Baidakova. 2020. Practice of Efficient Data Collection via Crowdsourcing: Aggregation, Incremental Relabelling, and Pricing. In *Proceedings of the 13th International Conference on Web Search and Data Mining*. 873–876.

[10] Axel G Ekström, Diederick C Niehorster, and Erik J Olsson. 2022. Self-imposed filter bubbles: Selective attention and exposure in online search. *Computers in Human Behavior Reports* 7 (2022), 100226.

[11] Nicole B Ellison, Charles Steinfield, and Cliff Lampe. 2007. The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. *Journal of computer-mediated communication* 12, 4 (2007), 1143–1168.

[12] Nicole B Ellison, Charles Steinfield, and Cliff Lampe. 2011. Connection strategies: Social capital implications of Facebook-enabled communication practices. *New media & society* 13, 6 (2011), 873–892.

[13] Nicole B Ellison, Jessica Vitak, Rebecca Gray, and Cliff Lampe. 2014. Cultivating social resources on social network sites: Facebook relationship maintenance behaviors and their role in social capital processes. *Journal of Computer-Mediated Communication* 19, 4 (2014), 855–870.

[14] Anna Fanoberova and Hanna Kuczkowska. 2016. Effects of source credibility and information quality on attitudes and purchase intentions of apparel products: A quantitative study of online shopping among consumers in Sweden.

[15] Claudia Flores-Saviaga and Saiph Savage. 2021. Fighting disaster misinformation in Latin America: the# 19S Mexican earthquake case study. *Personal and Ubiquitous Computing* 25, 2 (2021), 353–373.

[16] Brian J Fogg, Cathy Soohoo, David R Danielson, Leslie Marable, Julianne Stanford, and Ellen R Tauber. 2003. How do users evaluate the credibility of Web sites? A study with over 2,500 participants. In *Proceedings of the 2003 conference on Designing for user experiences*. 1–15.

[17] Ujwal Gadiraju, Ricardo Kawase, Stefan Dietze, and Gianluca Demartini. 2015. Understanding malicious behavior in crowdsourcing platforms: The case of online surveys. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 1631–1640.

[18] Eun Go, Eun Hwa Jung, and Mu Wu. 2014. The effects of source cues on online news perception. *Computers in Human Behavior* 38 (2014), 358–367.

[19] Angela R Gover, Shannon B Harper, and Lynn Langton. 2020. Anti-Asian hate crime during the COVID-19 pandemic: Exploring the reproduction of inequality. *American journal of criminal justice* 45 (2020), 647–667.

[20] Melissa A Hardy. 1993. *Regression with dummy variables*. Vol. 93. Sage.

[21] E. K. Hämäläinen, C. Kiili, M. Marttunen, E. Rikknen, and Pht Leppänen. 2020. Promoting sixth graders' credibility Evaluation of Web pages: An intervention study. *Computers in Human Behavior* (2020), 106372.

[22] Thomas J Johnson and Barbara K Kaye. 1998. Cruising is believing?: Comparing Internet and traditional sources on media credibility measures. *Journalism & Mass Communication Quarterly* 75, 2 (1998), 325–340.

[23] Thomas J Johnson and Barbara K Kaye. 2016. Some like it lots: The influence of interactivity and reliance on credibility. *Computers in Human Behavior* 61 (2016), 136–145.

[24] Nicholas A Jones and Michael Bentley. 2017. Overview of the 2015 national content test analysis report on race & ethnicity. *US Census Bureau, Suitland-Silver Hill, MD* (2017).

[25] Minjeong Kang. 2010. Measuring social media credibility: A study on a measure of blog credibility. *Institute for Public Relations* (2010), 59–68.

[26] Jinyoung Kim and Andrew Gambino. 2016. Do we trust the crowd or information system? Effects of personalization and bandwagon cues on users' attitudes and behavioral intentions toward a restaurant recommendation website. *Computers in Human Behavior* 65 (2016), 369–379.

[27] Lars König and Regina Jucks. 2019. Hot topics in science communication: Aggressive language decreases trustworthiness and credibility in scientific debates. *Public understanding of science* 28, 4 (2019), 401–416.

[28] Lars König and Regina Jucks. 2020. Effects of Positive Language and Profession on Trustworthiness and Credibility in Online Health Advice: Experimental Study. *Journal of medical Internet research* 22, 3 (2020).

[29] Lars König, Regina Jucks, et al. 2019. Influence of enthusiastic language on the credibility of health information and the trustworthiness of science communicators: Insights from a between-subject web-based experiment. *Interactive Journal of Medical Research* 8, 3 (2019), e13619.

[30] Junhao Li, Ville Paananen, Sharadhi Alape Suryanarayana, Eetu Huusko, Miikka Kuutila, Mika Mäntylä, and Simo Hosio. 2023. It is an online platform and not the real world, I don't care much: Investigating Twitter Profile Credibility With an Online Machine Learning-Based Tool. In *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval*. 117–127.

[31] Isabelle Lid Rosenholm. 2022. Health Communication on TikTok: A Qualitative Study of Credibility on A Humorous Platform.

[32] Mufan Luo, Jeffrey T Hancock, and David M Markowitz. 2022. Credibility perceptions and detection accuracy of fake news headlines on social media: Effects of truth-bias and endorsement cues. *Communication Research* 49, 2 (2022), 171–195.

[33] Debbie S Ma, Joshua Correll, and Bernd Wittenbrink. 2015. The Chicago face database: A free stimulus set of faces and norming data. *Behavior research methods* 47, 4 (2015), 1122–1135.

[34] Mark MacCarthy. 2020. Transparency requirements for digital social media platforms: Recommendations for policy makers and industry. *Transatlantic Working Group* (2020).

[35] Scott R Maier, Marcus Mayorga, and Paul Slovic. 2017. Personalized news stories affect men as well as women. *Newspaper Research Journal* 38, 2 (2017), 172–186.

[36] Scott R Maier, Paul Slovic, and Marcus Mayorga. 2017. Reader reaction to news of mass suffering: Assessing the influence of story form and emotional response. *Journalism* 18, 8 (2017), 1011–1029.

[37] Kelly Mathews, Jessica Phelan, Nicholas A Jones, Sarah Konya, Rachel Marks, Beverly M Pratt, Julia Coombs, and Michael Bentley. 2015. National Content Test: Race and ethnicity analysis report. *US Department of Commerce, Economics and Statistics Administration, US Census Bureau* (2015).

[38] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–23.

[39] Davood Mehrabi, Musa Abu Hassan, and Muhamad Sham Shahkat Ali. 2009. News media credibility of the Internet and television. *European journal of social sciences* 11, 1 (2009), 136–148.

[40] Miriam J Metzger and Andrew J Flanagin. 2015. Psychological approaches to credibility assessment online. *The handbook of the psychology of communication technology* 32 (2015), 445–466.

[41] Jonas Oppenlaender, Kristy Milland, Aku Visuri, Panos Ipeirotis, and Simo Hosio. 2020. Creativity on paid crowdsourcing platforms. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[42] Zizi Papacharissi. 2009. The virtual geographies of social networks: a comparative analysis of Facebook, LinkedIn and ASmallWorld. *New media & society* 11, 1-2 (2009), 199–220.

[43] Eyal Peer, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. 2017. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology* 70 (2017), 153–163.

[44] Richard E Petty, John T Cacioppo, and Rachel Goldman. 1981. Personal involvement as a determinant of argument-based persuasion. *Journal of personality and*

social psychology 41, 5 (1981), 847.

[45] Kevin Roitero, Michael Soprano, Beatrice Portelli, Massimiliano De Luise, Damiano Spina, Vincenzo Della Mea, Giuseppe Serra, Stefano Mizzaro, and Gianluca Demartini. 2021. Can the crowd judge truthfulness? A longitudinal study on recent misinformation about COVID-19. *Personal and Ubiquitous Computing* (2021), 1–31.

[46] Madison Elizabeth Sauls. 2018. *Perceived Credibility of Information on Internet Health Forums*. Ph. D. Dissertation. Clemson University.

[47] Laura Sbaffi and Jennifer Rowley. 2017. Trust and credibility in web-based health information: a review and agenda for future research. *Journal of medical Internet research* 19, 6 (2017), e218.

[48] Charles C Self. 2014. Credibility. In *An integrated approach to communication theory and research*. Routledge, 449–470.

[49] Patric R Spence, Kenneth A Lachlan, David Westerman, and Stephen A Spates. 2013. Where the gates matter less: Ethnicity and perceived source credibility in social media health messages. *Howard Journal of Communications* 24, 1 (2013), 1–16.

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

[51] C Strijbos, M Schluck, J Bisschop, T Bui, I De Jong, M Van Leeuwen, M von Tottleben, and SG van Breda. 2016. Consumer awareness and credibility factors of health claims on innovative meat products in a cross-sectional population study in the Netherlands. *Food Quality and Preference* 54 (2016), 13–22.

[52] Cass R Sunstein. 2006. *Infotopia: How many minds produce knowledge*. Oxford University Press.

[53] Kyle A Thomas and Scott Clifford. 2017. Validity and Mechanical Turk: An assessment of exclusion methods and interactive experiments. *Computers in Human Behavior* 77 (2017), 184–197.

[54] Sonja Utz. 2015. The function of self-disclosure on social network sites: Not only intimate, but also positive and entertaining self-disclosures increase the feeling of connection. *Computers in Human Behavior* 45 (2015), 1–10.

[55] Hans CM Van Trijp and Ivo A Van der Lans. 2007. Consumer perceptions of nutrition and health claims. *Appetite* 48, 3 (2007), 305–324.

[56] Amy Beth Warriner, Victor Kuperman, and Marc Brysbaert. 2013. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior research methods* 45, 4 (2013), 1191–1207.

[57] Senuri Wijenayake, Danula Hettiachchi, Simo Johannes Hosio, Vassilis Kostakos, and Jorge Goncalves. 2020. Effect of Conformity on Perceived Trustworthiness of News in Social Media. *IEEE Internet Computing* (2020).

[58] Dmitri Williams. 2006. On and off the'Net: Scales for social capital in an online era. *Journal of computer-mediated communication* 11, 2 (2006), 593–628.

[59] Anja Wölker and Thomas E Powell. 2018. Algorithms in the newsroom? News readers' perceived credibility and selection of automated journalism. *Journalism* (2018), 1464884918757072.

[60] Min Xiao, Rang Wang, and Sylvia Chan-Olmsted. 2018. Factors affecting YouTube influencer marketing credibility: a heuristic-systematic model. *Journal of Media Business Studies* 15 (07 2018), 1–26. https://doi.org/10.1080/16522354.2018.1501146

[61] Jia Zhou, Honglian Xiang, and Bingjun Xie. 2022. Better safe than sorry: a study on older adults' credibility judgments and spreading of health misinformation. *Universal Access in the Information Society* (2022), 1–10.