

Sharing, commenting, and reacting to Danish misinformation

A case study of cognitive attraction on Facebook

PETRA DE PLACE BAK^I & ETHAN WEED^{II}

^I DEPARTMENT OF MEDIA AND JOURNALISM STUDIES, AARHUS UNIVERSITY, DENMARK

^{II} DEPARTMENT OF LINGUISTICS, COGNITIVE SCIENCE AND SEMIOTICS, AARHUS UNIVERSITY, DENMARK

ABSTRACT

Social media facilitate a competition for users' limited attention by bringing various content together, from health advice to entertainment, and from updates from loved ones to misinformation. Especially misinformation has raised societal concern. We evaluated the influence of visual material and cognitive factors of attraction, specifically valenced sentiment, threat-related, intergroup-related, and social information, on engagement scores (i.e., shares, comments, and reactions). We analysed 356 misleading Danish Facebook posts sampled through the fact-checking association TjekDet's "entirely or partly false" web page by fitting a Bayesian zero-inflated negative binomial regression model. The study showed that videos and images were exceptionally strong predictors of engagement, especially shares. Positivity, negativity, and intergroup-related information also increased engagement, but social information and threat-related information reduced it. Our findings suggest that in a highly competitive online environment, some content biases are stronger than others. Finally, we discuss the potential moderators of their effect such as the users' reputation management strategies.

KEYWORDS: content biases, online engagement behaviour, social media communication, social transmission

Introduction

Social media platforms play a prominent role in modern communication, with adults now averaging nearly 2.5 hours of daily use (Statista, 2024). Among the changes brought about by the digital transformation of communication technologies are the removal of editorial filters and the extension of the potential reach of the users' content. This has resulted in an information abundance as various previously separated social contexts have come together (Marwick & boyd, 2011). Societal stakeholders, such as the World Health Organization, have expressed concern about the quality of information that might make reliable information harder to identify (WHO, 2020). Since the widespread adoption of social media platforms, digital misinformation has received increased mainstream and scholarly attention in recent years (for reviews, see Abu Arqoub et al., 2020; Bak et al., 2022).

Recommender algorithms play a key role in organising the vast amount of information on social media, but users' engagement behaviours – such as what they react to, comment on, and share – also significantly influence content visibility. This raises the question of which factors determine what attracts attention and drives engagement. Acerbi (2019) suggested that misinformation spreads on social media because it aligns with certain human cognitive preferences. Building on this idea, we explore the literature on cultural evolution and content biases in human communication. In this context, content biases refer to content types that are more attention-grabbing, memorable, and/or likely to be shared (Stubbersfield, 2022). We aim to test the hypothesis that information that taps into the cognitive preferences for certain types of content can boost online engagement.

The continuous registration of user data on social media makes it possible to test the relationship between such factors of attraction and the decision to transmit or engage with information with high ecological validity. So far, studies on content biases in online information-sharing have primarily focused on the relationship between sentiment (i.e., negativity and positivity bias) and the number of shares (see, e.g., De León & Trilling, 2021; Schöne et al., 2021; Youngblood et al., 2023). However, earlier experimental studies have found evidence of a bias toward social information (e.g., McGuigan & Cubillo, 2013; Stubbersfield et al., 2015) and threat-related information (e.g., Blaine & Boyer, 2018; Nairne et al., 2007), among others. While a few studies have examined the frequency of factors of attraction beyond sentiment in popular content such as urban legends (Stubbersfield et al., 2017) and online misinformation (Acerbi, 2019), these studies did not evaluate the relationship between the presence of factors of attraction and engagement scores (i.e., shares, reactions, and comments).

Based on the literature on content biases, we chose to focus on the following factors: positivity, negativity, social information (i.e., information about third-party interactions), threat-related information, and intergroup-related information (for reviews, see Stubbersfield, 2022, 2025). Empirically, we drew on the case of misinformation because it is societally relevant and theoretically significant since

misinformation spreads despite being maladaptive and potentially harmful, for example, if users adopt the wrong health behaviours. Further, misinformation may benefit from being malleable to be optimally attractive, whereas authentic information is limited to the truth. We addressed the following research questions:

RQ1. To what extent are factors of attraction (i.e., negativity, positivity, social information, and threat-related and intergroup-related information) present in misleading Danish Facebook posts?

RQ2. How does the presence of factors of attraction and visual material influence the success of misleading Danish Facebook posts?

We analysed a dataset of Danish Facebook posts gathered from the Danish fact-checking agency TjekDet's website. A considerable number of posts are associated with the Covid-19 pandemic, and therefore we chose to test the potential influence of this topic on engagement. Finally, since the layout of Facebook posts vary greatly depending on whether they include visual material, we also tested the effect of including an image or video.

Danish misinformation on Facebook

In the study, we drew on the empirical case of misleading Facebook posts in the Danish context. In line with earlier work, we embraced an empirically grounded interpretation of misinformation, characterising misinformation as content identified as “partly or entirely false” by accredited fact-checking organisations (Nissen et al., 2022). Empirical contributions to understand and classify misinformation have primarily focused on English data from Twitter (Bak et al., 2022). On Twitter, a case study from the 2019 Danish national election suggested that misinformation is not common in the Danish context compared with the UK (Derczynski et al., 2019). However, compared with Twitter, Facebook is a much more commonly used social media platform, both globally (Statista, 2023) and in Denmark (DR Analyse, 2024). In 2023, six out of ten adults (15–75 years old) used Facebook daily, with some variations between age groups, according to Danmarks Radio's media development report (DR Analyse, 2024).

Like other Nordic countries, Denmark has a high Internet penetration rate of 98 per cent, according to the *Reuters Institute Digital News Report* (Newman et al., 2024). While news consumption patterns in the Nordic countries are generally similar compared with other regions (Newman et al., 2024), Denmark stands out. In fact, 19 per cent of Danish respondents reported social media as their main gateway to online news, compared with just 10 per cent in Norway, 11 per cent in Sweden, and 12 per cent in Finland (Newman et al., 2024). Furthermore, only 17 per cent of Danish respondents were willing to pay for online news, far lower than the 40 per cent of Norwegians, 31 per cent of Swedes, and 20 per cent of Finns. This suggests that, among the Nordic countries, Danes rely more heavily on social media for news and less on traditional news outlets, potentially making them more vulnerable to misinformation than their Nordic neighbours.

Factors of attraction and online engagement

Research has shown that certain types of information are more likely to be noticed, memorised, and retransmitted than others (Stubbersfield, 2025), so-called transmission biases. These can be either content biases (e.g., negativity bias) or context biases (e.g., learning from prestigious or successful individuals) (Kendal et al., 2018). Apart from biases toward emotion-expressing content, empirical contributions to the study of content biases in naturally occurring human communication have primarily focused on their frequency in popular content types; for example, Stubbersfield and colleagues (2017) found factors of attraction in 92 per cent of a sample of popular urban legends, primarily emotional content. Similarly, Acerbi (2019) found that 86 per cent of a sample of fake news headlines contained elements that could be traced back to factors of attraction. In line with these findings, we expect the factors of attraction to be present in a large proportion of the misleading Facebook posts.

There is a growing number of studies that have modelled the relationship between the expression of sentiment or discrete emotions and online sharing, and a limited number of studies that have looked at intergroup-related biases (e.g., Rathje et al., 2024). To form testable hypotheses, we reviewed the main literature in relation to the factors included in the study – that is, negativity, positivity, social information, threat-related information, intergroup-related information, and visual material.

Negativity bias is well-documented within psychology (Rozin & Royzman, 2001). It has an intuitive evolutionary rationale: It is better to expect the worst and be positively surprised than the other way around. This way, sensitivity to negative information may increase the chance of survival. Transmission chain studies have revealed negativity to be more transmitted than positivity (Bebbington et al., 2017) and that ambiguous stories are rendered more negatively over repeated transmissions. An empirical study of two centuries of Anglophone fiction found a general decline in the use of emotional words, driven by a decrease in positive language (Morin & Acerbi, 2017).

Berger and Milkman (2010) were the first to test the effect of sentiment and emotion expressions on the sharing of content in digital media. Drawing on a sample of *New York Times* articles, the study found that emotion-expressing content is more likely to reach the most e-mailed list than neutral content. While the effect was strongest for positivity compared with negativity, anger turned out to be the strongest predictor among the discrete emotions, suggesting that emotions of the same valence can have different effects. More recent work has focused on emotions and information sharing on social media. Youngblood and colleagues (2023) found a relationship between negativity in voter-fraud conspiracy theories on Twitter and a higher number of retweets. In contrast, Berriche and Altay (2020) found a positivity bias on Facebook and that phatic posts – that is, small talk such as “how are you?” – were the strongest predictor of engagement. Others have found that positivity increases the spread of verified accurate content (King & Wang, 2023; Luo et al., 2023), but not misinformation.

A reason why emotion-expressing content receives higher engagement levels could be because it captures attention (Brady et al., 2020). Based on the reviewed literature, our first hypothesis is as follows:

H1: Both positive and negative information will receive more engagement than neutral posts because it captures attention better, hereby increasing the chances of engagement.

Another content type that has received attention in transmission chain experiments is social information, such as gossip and information about interactions between third parties (e.g., McGuigan & Cubillo, 2013; Stubbersfield et al., 2015). In freely forming conversation, social information makes up approximately two-thirds of the topics discussed (Dunbar, 2004). The social brain hypothesis holds that the human brain evolved to live in complex social groups (Dunbar, 2009), for example, to support corporation and communication. Experimental findings suggest that both children (McGuigan & Cubillo, 2013) and adult (Mesoudi et al., 2006) participants are more likely to transmit social than non-social information. In addition, a study on urban legends found that social information is better recalled than non-social information (Stubbersfield et al., 2015). Finally, in the digital context, Acerbi (2019) found social information and celebrity information to be the most frequently present factors of attraction in a dataset of popular fake news headlines. Similarly, Berriche and Altay (2020) found half of their sample of Facebook posts to be about social information, but social information did not increase engagement. Instead, the main driver of engagement was phatic posts – that is, posts aimed at establishing a social connection. Based on the literature, we pose our second hypothesis:

H2: Social information will receive more engagement than non-social information, in line with the social brain hypothesis and because social media platforms are primarily used for exchanging social information.

Threat-related information concerns anything related to human survival, such as well-being and health. In an experimental study of the transmission of rumours, Blaine and Boyer (2018) found that threat-related information is transmitted more often than emotion-expressing content, even for threats that are very unlikely. Nairne and colleagues (2007) found survival-related material to be better remembered compared to control conditions, indicating that human memory has evolved to remember information relevant for survival. Similarly, Stubbersfield and colleagues (2015) found threat-related information to be better recalled than control material, and the combination of social and threat-related information to be significantly stronger than threat alone. A motivation to transmit threat-related information is that sources of threat-related information are evaluated as more competent (Boyer & Parren, 2015). Moreover, on Twitter, Mousavi and colleagues (2022) found that threats “to personal or societal core values of target audiences” increased the engagement. However, whether this generalises to a broader definition of threat-related information is currently unknown. We thus pose our third hypothesis:

H3: Threat-related information will increase the overall engagement score because it is adaptive to pay attention to threats and, being the source of threat-related information, can be associated with reputational benefits.

We also focused on intergroup-related information. Intergroup-related information differs from social information in that it describes relationships at the group level, for example, religious or ethnic groups, rather than information about identifiable individuals. It taps into intergroup bias, that is, the tendency to favour the group one identifies with (ingroup) over the outgroup (Hewstone et al., 2002). On Twitter, users tend to speak positively about their political ingroup and negatively about their outgroup (Stieglitz & Dang-Xuan, 2013). It remains unclear how intergroup-related information influences spread on social media, but information that aligns with preexisting social categories is selected for (Lyons & Kashima, 2001; Martin et al., 2017). Rathje and colleagues (2024) found posts about a political outgroup to be a stronger positive predictor of shares on Facebook and Twitter than the more established predictors of negativity and moral-emotional language. Our fourth hypothesis is thus as follows:

H4: Intergroup-related information will lead to higher engagement scores because it serves as a strategy for signaling group membership.

Finally, we consider the inclusion of visual material. The cognitive system has a longer history of engaging with visual material rather than text, for example, when our ancestors scanned their ecological surroundings. Visual material also differs from text because it stands out by being colourful and, in the case of video, includes motion, which might be efficient attention-grabbers in a crowded information space. Research from the field of advertising suggest that ads with visual elements grab attention better than purely textual ads (Goodrich, 2011; Pieters & Wedel, 2004). On Twitter, Bruni and colleagues (2012) found tweets with links to videos or images to be more retweeted than posts without multimedia content. In a more recent study, Li and Xie (2020) found that including an image in a tweet increased engagement. We thus propose our fifth hypothesis:

H5: The inclusion of visual material, that is, images and videos, leads to higher engagement scores because it grabs attention better than purely text-based posts.

Data and methodology

Data collection

The dataset was collected via the EFSCN and IFCN-certified fact-checking organisation TjekDet's [CheckIt] website, where the fact-checkers exemplify claims rated as "entirely or partly false" through direct links to Facebook. Fact-checking agencies are a common data source in studies of misinformation due to the convenience of an already annotated sample (López-García et al., 2021; Song et al., 2021; Vosoughi et al., 2018). The data were collected by scraping

all external links with the word “Facebook” from the “entirely or partly false” web page.

The described data-sampling method resulted in a list of links that were then manually visited by the first author to extract data about the posts to be used in the analysis. For each post, we collected the text, visual material, and the year the post was published on Facebook, as well as the engagement score (i.e., total number of reactions, shares, and comments). This approach resulted in a dataset of 356 Facebook posts, after four were excluded because they appeared more than once. The misleading claims were posted on Facebook during 2011–2023, with most entries being from 2021 (110 posts) and 2022 (95 posts). The dataset is considered sufficient in size for the purpose of this study (for similar samples, see Acerbi, 2019; Berriche & Altay, 2020), as the study relies on the idea that individual-level interactions result in population-level aggregates (Acerbi, 2016).

Defining dependent variables

We chose engagement scores as an operational measure of online information sharing. On Facebook, users engage with posts through comments, shares, and reaction functions, such as “like”, “love”, and “wow”. Our analysis scrutinises these engagement types collectively and individually, exploring potential disparities in user behaviours to enrich our understanding of the dynamics surrounding misinformation dissemination. Engagement scores are well-established as measures of spread in the context of social media, especially in studies quantifying the effect of sentiment expressions (e.g., Berriche & Altay, 2020; Ferrara & Yang, 2015; Fine & Hunt, 2023; Schöne et al., 2021).

Defining independent variables

The data was manually coded for the presence or absence of the content factors. The posts were rated as positive, negative, or neither, thus treating positive and negative sentiment as mutually exclusive categories. Hence, each post could contain a maximum of four factors of attraction. The categories are defined according to the descriptions below (detailed instruction for coders can be found in Supplementary Material A).

Table 1 Categories and descriptions

Category	Description
Social	The content concerns intense and noticeable social relationships (e.g., gossip, cheating, group alliances, controversies) but also everyday interactions and relationships (e.g., friends, family).
Threat-related	The content concerns possible threats (e.g., illness, violent acts, dangerous situations, death, and involuntary abortion).
Positive	The content expresses emotions that can be considered overall as positive (e.g., amusement, love, joy, etc.).
Negative	The content expresses emotions that can be considered overall as negative (e.g., sadness, regret, fear, anger, etc.).
Intergroup-related	The content concerns intergroup relationships, differences, or attitudes based on group membership (e.g., based in cultural, religious, social, and value differences).

Comments: The definitions are inspired by earlier work (see Acerbi, 2019; Berriche & Altay, 2020; Scheffer et al., 2021; Stubbersfield et al., 2017).

The first author and a trained assistant independently coded 290 posts based on the absence or presence of factors of attraction from the five categories detailed in Table 1. To assess the reliability of the coding, the agreement was measured using Cohen’s kappa (κ), a measurement adjusted for chance agreement often used in social and behavioural sciences to calculate intercoder reliability (Warrens, 2015). The agreement resulted in $\kappa = 0.669$, which is interpreted by Cohen as substantial agreement (see Supplementary Material B for category-specific intercoder reliability scores). All points of disagreement were discussed to reach consensus. In addition to the factors of attraction, the following variables were coded by the main author alone, as these do not require interpretation: emojis, links, images, videos, and topic (i.e., Covid and non-Covid).

Concerning data access, the annotations and engagement scores are made available (OSF, 2023) together with the code used for scraping the external links from TjekDet and implementing the statistical analysis described below. This allows for repetition of the statistical analyses. To secure GDPR compliance, the Facebook posts cannot be shared. To convey the idea of how the posts look and were analysed, we have created the following mock example with an image generated by ChatGPT Plus:

Figure 1 Mock example of how the posts might look



Comments: The image of cars included in the mock post was generated by ChatGPT Plus.

This post would be rated as negative in tone due to the words “scam” and “lying” in combination with double use of exclamation marks. It also contains intergroup-related information, as the “politicians” are mentioned as lying and implicitly also responsible for the “scam”. It does not contain any explicit threats or social information, nor does it convey positive sentiment.

Statistical analysis

We evaluated the relationship between the coded predictors and the level of engagement by fitting a Bayesian zero-inflated negative binomial regression model. All predictors were modelled as fixed effects. Given the exploratory nature of the study, we chose conservative priors for our models that would constrain the models within a reasonable range (e.g., the models will not predict negative numbers of shares), without exerting any undue influence on the results. For the model intercept, we used a weakly informative prior, a student *t*-distribution, with three degrees of freedom, a location of 3.5, and scale of 2.8. For the shape parameter of the model, we used an inverse gamma distribution with shape and

scale parameters of 0.04 and 0.3, respectively. This ensures the model will not predict numbers of shares less than zero, and gives a good approximation of the over-dispersed tendencies (Greene, 2003; Hilbe, 2011) typical of social media engagement scores (e.g., Apenteng et al., 2020; Gross & Von Wangenheim, 2022), in which the vast majority of posts receive very little engagement, but a small number of posts have extremely high levels of engagement.

All analyses were performed in R (R Core Team, 2023). Data were modelled using the *brms* package (Bürkner, 2017). Graphical posterior predictive checks were performed using the *bayesplot* package (Gabry & Mahr, 2022), model fit was further evaluated with the *DHARMA* package (Hartig, 2024), and Bayesian R^2 was calculated using the method described by Gelman and colleagues (2019) and implemented in the *performance* package (Lüdtke et al., 2021). To make our results more easily interpretable, we converted the model coefficient estimates to incidence rate ratios by exponentiating the estimates.

Results

Of the 356 Facebook posts, the majority (326) contained one or more cognitive attractive factors (92%). Of these, 114 (32%) contained two cognitive attractive factors, 101 (28%) contained three, 83 (23%) contained one, and 28 contained four (8%). 30 (8%) contained none

Analysis of descriptive statistics

Figure 2 shows a difference at the topic level the proportion of posts containing each of the cognitive factors of attraction when the information is related to Covid-19 or not. In general, threat-related information, negativity, and intergroup-related information are the most frequently occurring factors of attraction. Intuitively, threat-related information and negativity are more present in the posts related to Covid-19 than those that are not.

Figure 2 Proportion of factors by Covid status (per cent)

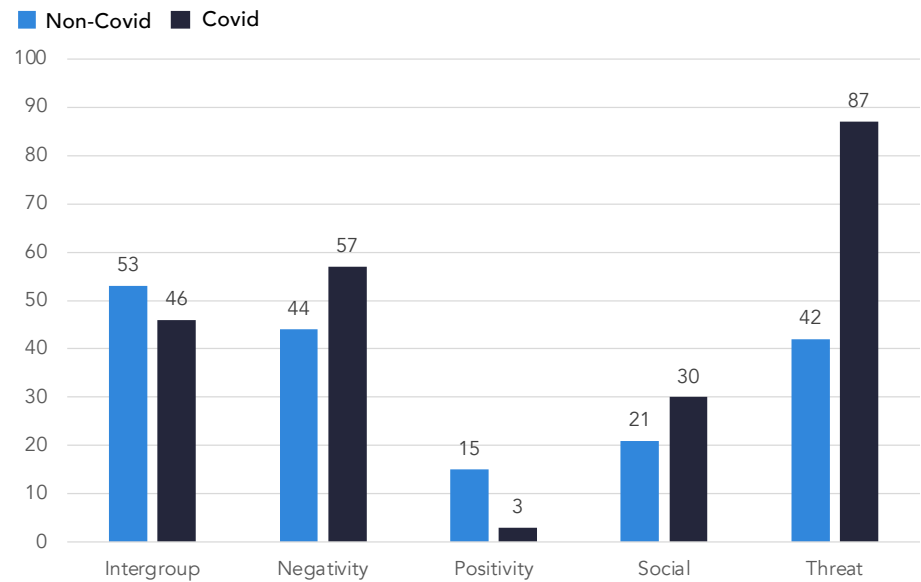


Figure 3 shows the co-occurrences between the different factors of attraction. Most notably, social information often co-occurs with threat-related information (69%). Similarly, intergroup-related information often co-occurs with threats (79%) and negative information (64%). Threats are often formulated in a negative tone (58%).

Figure 3 Co-occurrences of factors of attraction (row baseline, per cent)

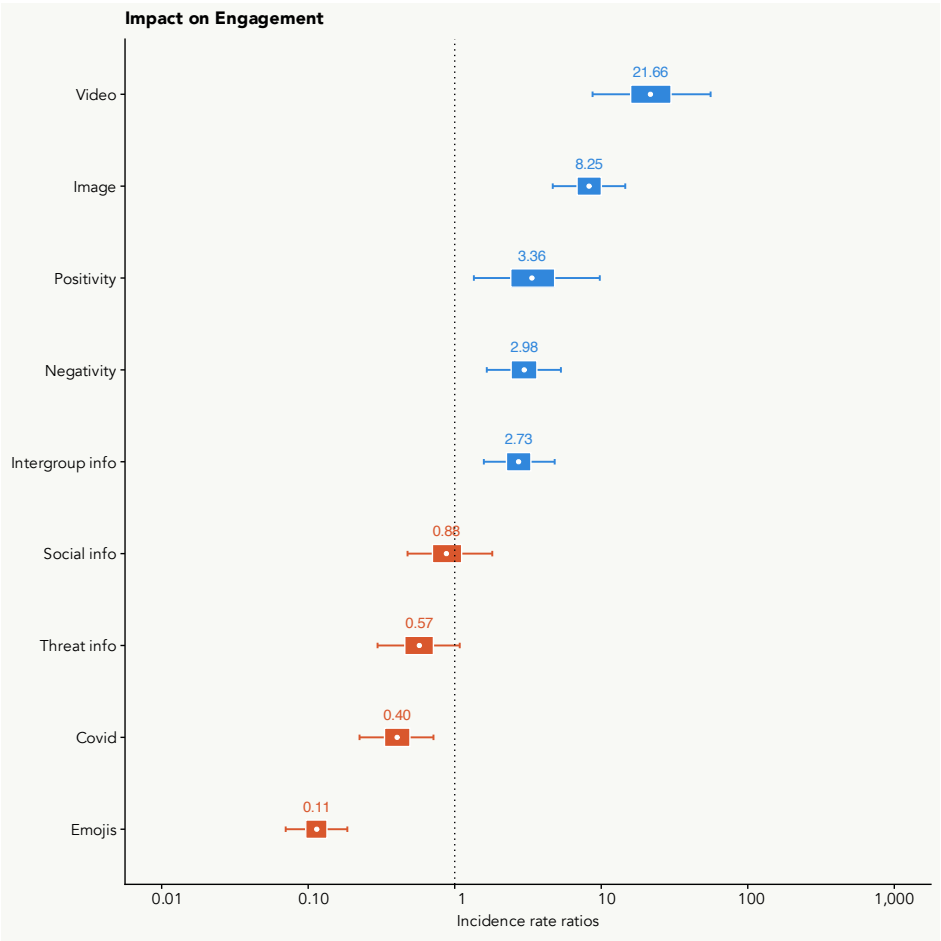
	Social	Threat	Positivity	Negativity	Intergroup	Row baseline
Social	100%	69%	6%	60%	51%	100%
Threat	26%	100%	4%	58%	56%	75%
Positivity	21%	36%	100%	0%	29%	50%
Negativity	30%	78%	0%	100%	60%	25%
Intergroup	28%	79%	5%	64%	100%	0%

Comments: The row is the baseline, which means that, for example, of the times that intergroup-related information is present, it co-occurs with social information 28 per cent of the time, threat-related information 79 per cent of the time, and so on, whereas of the times social information occurs, it co-occurs with intergroup-related information 51 per cent of the time. The difference in percentage is caused by a difference in frequency between the categories.

Analysis of effects on engagement scores

The modelled predictors accounted for 6 per cent of the variance in engagement (Bayesian $R^2 = 0.063$ [95% CI {0.015, 0.167}]). Video, image, emotional valence, and intergroup information have the largest influence on the estimated change in predicted engagement, while the other variables are less influential (see Figure 4).

Figure 4 Incidence rate ratios of predictor variables on total engagement



Comments: Values lower than 1 have a negative influence on engagement, while values above 1 have a positive influence. Dots indicate the median of the posterior distributions, as do the values above each point. Whiskers indicate 95 per cent credible intervals, and bars indicate 50 per cent credible intervals.

Looking at Figure 4, the overall finding is that visual material (i.e., video and images) is the strongest predictor of increased engagement. Further, positivity, negativity, and intergroup-related information is also associated with increased engagement. Social information has a negative effect, and threat-related information, Covid-19, and the inclusion of emojis also reduce the overall engagement.

Table 2 shows the incidence rate ratios of total engagement and the three engagement types – comments, shares, and reactions – individually. The most noticeable aspect across engagement types is that the presence of videos is a far stronger predictor (numbers rounded) of shares ($\times 52.52$) than comments ($\times 3.19$) and reactions ($\times 3.02$). The same pattern, although less pronounced, is shown for the other visual element, images, which increases the number of predicted shares by 7 times, comments by 2 times, and reactions by 3 times. Secondly, negatively valenced information influences shares ($\times 4.46$) much more than comments ($\times 2.28$) and reactions ($\times 1.65$). Further, intergroup-related information resulted in 3.86 times the number of shares and 2.54 times the number of reactions but had no influence on comments ($\times 0.99$). The remaining predictor variables show variance across forms of engagement.

Table 2 Incidence rate ratios for total and subtypes of engagement (i.e., comments, shares, and reactions)

Predictors	Total	Comments	Shares	Reactions
Video	21.66 (14.07)	3.19 (1.91)	52.52 (43.2)	3.02 (1.69)
Image	8.25 (3.18)	2.8 (1.25)	12.54 (7.44)	3.72 (1.44)
Positivity	3.36 (2.4)	1.85 (1.34)	2.76 (2.85)	2.51 (1.6)
Negativity	2.98 (1.21)	2.28 (0.92)	4.46 (2.32)	1.65 (0.6)
Intergroup info	2.73 (1.06)	0.99 (0.39)	3.86 (1.86)	2.54 (0.88)
Social info	0.88 (0.41)	1.08 (0.48)	0.92 (0.55)	0.96 (0.37)
Threat info	0.57 (0.26)	0.82 (0.42)	0.87 (0.45)	0.65 (0.27)
Covid	0.4 (0.16)	0.15 (0.07)	0.38 (0.2)	0.36 (0.13)
Emojis	0.11 (0.04)	0.77 (0.3)	0.05 (0.02)	0.22 (0.07)

Comments: The incident rate ratios shows how many times more or less likely a post is to get engagement per predictor variable; the closer to 1, the less influence the predictor has on engagement. Values lower than 1 have a negative influence on engagement, while values above 1 have a positive influence. The number in parentheses is the interquartile range.

Discussion

We tested the influence of factors of attraction and visual material on the overall number of engagements (shares, reactions, and comments) that Danish misinformation received. The study showed a high frequency of the factors of attraction in the sample of misleading posts, with 92 per cent containing one or more factors. The high proportion of content that taps into content biases align with previous findings (e.g., Acerbi, 2019; Berriche & Altay, 2020; Stubbersfield et al., 2017). Hypotheses 1, 4, and 5 were all supported, as visual material (video

and images), sentiment-expressing, and intergroup-related information received higher engagement levels. In contrast, hypotheses 2 and 3 were not supported, as social and threat-related information did not lead to higher engagement. This indicates that a high frequency of factors of attraction does not necessarily translate into high engagement.

Consistent with prior research (e.g., Bruni et al., 2012; Goodrich, 2011; Li & Xie, 2020; Pieters & Wedel, 2004), our findings indicate that visual material results in higher engagement levels compared with non-visual content. This outcome is expected, given that visual content is particularly effective in capturing attention amidst the vast array of information on social media. Importantly, this effect appears to be amplified in the context of misinformation. Visual material can be perceived as corroborative evidence, thereby reducing the perceived reputational risks associated with sharing potentially false claims (Altay et al., 2022). In essence, visuals may serve as a heuristic for validation, encouraging users to share content they might otherwise hesitate to disseminate. An example from the data is posts sharing the trailer for the conspiracy theorist “documentary”, *Plandemic: The Hidden Agenda Behind COVID-19* (Funke, 2020).

That threat-related and social information decreases the engagement level can be interpreted as an indication that these biases are primarily occurring in the memorisation and recall phase of transmission. Recall is not a requisite for online information-sharing where users can simply use the share function. This interpretation aligns with Acerbi’s (2021) experimental findings, in which the threat-related information was favoured only in the condition of mimicking oral transmission and not in the digital-sharing condition. Further, experimental work on urban legends has found social and threat-related information to be advantageous in the encode-and-retrieve phase (Stubbersfield et al., 2015), but not in the choose-to-receive and choose-to-transmit phases. The finding that threat-related information decreases the engagement score aligns with previous findings on Facebook (Berriche & Altay, 2020). A viable interpretation would be that the cognitive constraints on memorisation and reproduction in oral communication are what causes the bias toward threat-related information. Alternatively, or perhaps additionally, Covid-19 was a prominent theme in the data, which might have caused users to experience threat fatigue.

The affordances of social media – for example, that what we engage with is visible and persists across time and space (Bucher & Helmond, 2018) – may further moderate what we engage with. In the case of social information, it is not associated with an increase in engagement scores; however, it might still be attention-grabbing. Instead, users may abstain from engaging with it because of reputational concerns, such as not wanting to be caught gossiping and, in response, reserve such information for more private settings. Intergroup-related information, on the other hand, is less risky to share because it signals group membership and mitigates the risk of being caught gossiping. Future research could compare viewership data to engagement scores to test whether there are some types of content that have high numbers of views but little engagement. Alternatively, misleading social information might not spread because the wealth

of authentic social information available online overshadows such content in terms of salience. Finally, traditional social media platforms such as Facebook rely on social connection (Metzler & Garcia, 2023), in contrast to newer, more heavily algorithmic-driven platforms such as TikTok. It is possible that the goal of maintaining social relationships overrides the biases toward social and threat-related information. As for positivity and negativity bias, large-scale studies have found that expressions of valenced sentiment increase spread on social media consistently across several languages (e.g., Mousavi et al., 2022; Tsugawa & Ohsaki, 2015). This aligns well with previous findings that emotion is privileged in early attention capture (Brady et al., 2020). Thus, our study indicates that these previous findings also generalise to the case of Danish misinformation.

Finally, the analysis shows that the three most frequent factors were threat-related information, negatively valenced information, and intergroup-related information. A plausible reason why these three biases are dominant is that Covid-19 was a prominent theme in the dataset. A high proportion of the dataset conveyed threats, for example, warning against vaccination, and sought to attribute responsibility for the pandemic to the outgroup, for example, the political elite or ethnic minorities. However, the study also indicated that there is not necessarily a relationship between the frequency of a trait and its transmission via engagement. That said, frequency does signal that these are factors that misinformed users in their communication on social media. Thus, frequency can be interpreted as a different signal indicating a preference in the choose-to-transmit phase, in addition to engagement scores. Further, the discrepancy between the prevalence of certain factors and their impact on engagement can be interpreted in multiple ways. Users might be unaware that visual material, for instance, significantly boosts engagement. Alternatively, it may reflect the higher costs associated with producing visual content, such as time and technical skills. Additionally, users may have objectives beyond maximising engagement. For instance, they might prioritise sharing threat-related information quickly to alert their network and to be perceived as cooperative and valuable members to the group.

Engagement types as distinct digital behaviours

Factors of attraction influence interactions, comments, and shares differently. Analysing the engagement types individually (see Table 2) reveals that the primary effects are driven by changes in the engagement type shares, especially for visual material. The share category exhibits the greatest variance, with a median of 3 and a maximum of 406,000 shares for a single post (see Supplementary Material C, Table C2). A plausible explanation for why visual material and factors of attraction primarily influence sharing behaviour is that users may be reluctant to share information publicly if they doubt its veracity. However, when a post includes visual content, it may serve as additional support for the claim, making users more likely to consult others about the veracity of the claim. In such cases, the share function offers a strategic advantage, allowing users to consult trusted others by sharing in private chats, closed groups, or

on their own profile with restricted audience settings. In contrast, reacting or commenting on a post immediately exposes a user's stance – such as showing anger or agreement – to a large and undefined audience, and thus, potentially involves reputational damage. This suggests that reactions and comments may be preferred when users are more certain of the claim's truthfulness, either to support it or actively debunk it.

The observation that the types of engagement contribute differently to the overall score suggests that distinguishing different modes of engagement can provide insights into how various types of content appeal to specific behaviours. Compared to oral communication, the share function most closely resembles transmission. From a user perspective, reactions and comments are more subtle ways of extending a post's reach. However, both comments and reactions still signal to other users and the algorithm that the post merits engagement.

Interactions with additional factors

According to the theory of cultural evolution, all other things being equal, cognitively appealing content will spread more than equivalent non-appealing content. However, such conditions are only obtainable in controlled experimental settings. On social media, algorithmic content curation, context biases, and social factors are also at play. This may explain why the model implemented in the study only accounted for 6.3 per cent of the variance in engagement. While this may appear small, considering the scale of social transmission on digital social media, 6.3 per cent can make a large difference in what users see. While it is beyond the scope of this article to address in detail the moderating factors introduced on social media, we briefly discuss them to illustrate that the processes of cultural attraction do not happen in a vacuum. When discussing cultural transmission on social media, it is relevant to consider a broader context; thus, this study feeds into a wider research agenda on why some information spreads while other information vanishes.

Algorithmic content curation may cause the over-dispersed tendencies of social media engagement scores (e.g., Apenteng et al., 2020; Gross & Von Wangenheim, 2022), with a few viral posts and many more that go without engagement. A possible assumption in this regard is that algorithms are optimised for prolonging attention, for instance, by amplifying the reach or visibility of negatively valenced information. In a similar vein, algorithmic meddling can also limit the reach of certain content. However, understanding how algorithms work remains a challenge. Additional factors that interfere with the results may be users' perception of the audience and concerns for reputation management. In addition, context factors such as the number of followers or level of popularity may also influence what gets commented on, shared, and reacted to. That said, de Oliveira and Albuquerque (2021) found that social prestige does not determine spread of misinformation on Twitter. In short, the reasons for both sharing and engaging with online information are manifold, but cognitive factors of attraction seem to be part of the explanation.

Limitations and future directions

The first limitation this study faces is the relatively small sample size. However, this was also what made manual coding feasible. Future studies can adapt the methodological design and scale it through natural language processing tools; however, this would require training the natural language processing tools to categorise other cognitive factors of attraction besides sentiment and emotions. Further, whether the observed patterns transfer to other social media platforms and cultural contexts is a task for future research to unveil. However, as discussed, the findings align well with previous efforts in different national contexts. Finally, the timespan of the data stretches across several years, and it is possible that changes to algorithmic content curation in the same period accounts for some of the difference in engagement rates. An obvious extension of the study would be to test the same hypotheses on authentic information to see whether the patterns transfer across types.

Moreover, the size and sampling strategy of the dataset impose some limitations on the generalisability of our findings. TjekDet selects posts for fact-checking based on their general relevance to the public (TjekDet, 2020), which may result in a dataset that is not fully representative of misinformation, and even less so of general information. Nevertheless, our primary objective was to expand on studies examining the frequency of factors of attraction in popular online content (Acerbi, 2019; Stubbersfield et al., 2017) by evaluating their relationship with engagement metrics. Future research could test the robustness of our findings in larger and more diverse datasets, potentially by utilising advancements in language models for more comprehensive annotation.

Finally, studying cognitive factors of attraction in an ecologically valid setting is messy, as multiple factors other than the included factors exert influence on the decision to engage with information, for example, reputational concerns (Altay et al., 2022), and the entanglement of algorithmic and social drivers (Metzler & Garcia, 2022).

Despite these limitations, this article presents an interdisciplinary methodological framework combining cultural attraction theory and media studies and a fine-grained analysis of how factors of attraction scale in contemporary media systems. Additionally, the article adds to the field of misinformation studies by focusing on a platform other than Twitter, in an understudied cultural context (Bak et al., 2022), and by elaborating on the cultural success of misinformation (see also Acerbi, 2019; Berriche & Altay, 2020; Youngblood et al., 2023).

Conclusion

In this article, we have shown a high frequency of factors of attraction in a sample of misleading Facebook posts. Still, only visual material, positivity, negativity, and intergroup-related information are associated with higher engagement scores. The effect is primarily driven by the engagement type share, indicating that the different types of engagement – shares, comments, and reactions – mean different things. Overall, the model can explain 6.3 per cent of the variance in

engagement, suggesting that included factors of attraction do influence social media communication. However, the specific affordances of social media may moderate certain types of biases; for instance, intergroup-related information might be shared more than social information to avoid reputational damage from openly gossiping. Visual material has, unsurprisingly, the strongest influence on engagement scores, indicating that it is an efficient attention-grabber. A future step would be to augment the analysis to a larger dataset by training a classifier. This would allow us to determine whether the findings presented here generalise outside the context of misinformation on Danish Facebook.

Acknowledgements

We thank Professor Anja Bechmann for feedback on the manuscript, software developer Lea Riksted for support with scraping Facebook links, and student helper Luna Sol Korsholm Mathiesen for manual coding.

The corresponding author received 10,000 Danish kroner by PolitikenFonden to hire a student assistant to take part in the manual coding (grant no. 168/23).

Part of the dataset – that is, the annotations of presence of content biases and engagement scores – are available (OSF, 2023). The direct links, texts, and visual material are not shared. This is to protect the identity of the users, as some posts contain sensitive information such as health records.

The analysis scripts necessary for reproducing all analyses and figures are available (OSF 2023).

References

- Abu Arqoub, O., Elegu, A. A., Efe Özad, B., Dwikat, H., & Oloyede, F. A. (2020). Mapping the scholarship of fake news research: A systematic review. *Journalism Practice*, 16(1), 56–86. <https://doi.org/10.1080/17512786.2020.1805791>
- Acerbi, A. (2016). A cultural evolution approach to digital media. *Frontiers in Human Neuroscience*, 10. <https://doi.org/10.3389/fnhum.2016.00636>
- Acerbi, A. (2019). Cognitive attraction and online misinformation. *Palgrave Communications*, 5(1), 15. <https://doi.org/10.1057/s41599-019-0224-y>
- Acerbi, A. (2021). *From storytelling to Facebook: Content biases when retelling or sharing a story* [Preprint]. Open Science Framework. <https://doi.org/10.31219/osf.io/br56y>
- Altay, S., Hacquin, A.-S., & Mercier, H. (2022). Why do so few people share fake news? It hurts their reputation. *New Media & Society*, 24(6), 1303–1324. <https://doi.org/10.1177/1461444820969893>
- Apenteng, B. A., Ekpo, I. B., Mutiso, F. M., Akowuah, E. A., & Opoku, S. T. (2020). Examining the relationship between social media engagement and hospital revenue. *Health Marketing Quarterly*, 37(1), 10–21. <https://doi.org/10.1080/07359683.2020.1713575>
- Bak, P. de P., Walter, J. G., & Bechmann, A. (2022). Digital false information at scale in the European Union: Current state of research in various disciplines, and future directions. *New Media & Society*, 25(10), 2800–2819. <https://doi.org/10.1177/14614448221122146>
- Bebbington, K., MacLeod, C., Ellison, T. M., & Fay, N. (2017). The sky is falling: Evidence of a negativity bias in the social transmission of information. *Evolution and Human Behavior*, 38(1), 92–101. <https://doi.org/10.1016/j.evolhumbehav.2016.07.004>
- Berger, J., & Milkman, K. L. (2010). *Social transmission, emotion, and the virality of online content* [Marketing Science Institute Working Paper Series 2010, Report No. 10-114]. https://thearf-org-unified-admin.s3.amazonaws.com/MSI/2020/06/MSI_Report_10-114.pdf

- Berriche, M., & Altay, S. (2020). Internet users engage more with phatic posts than with health misinformation on Facebook. *Palgrave Communications*, 6(1), Article 71. <https://doi.org/10.1057/s41599-020-0452-1>
- Blaine, T., & Boyer, P. (2018). Origins of sinister rumors: A preference for threat-related material in the supply and demand of information. *Evolution and Human Behavior*, 39(1), 67–75. <https://doi.org/10.1016/j.evolhumbehav.2017.10.001>
- Boyer, P., & Parren, N. (2015). Threat-related information suggests competence: A possible factor in the spread of rumors. *PLOS ONE*, 10(6), e0128421. <https://doi.org/10.1371/journal.pone.0128421>
- Brady, W. J., Gantman, A. P., & Van Bavel, J. J. (2020). Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology: General*, 149(4), 746–756. <https://doi.org/10.1037/xge0000673>
- Bruni, L., Francalanci, C., & Giacomazzi, P. (2012). The role of multimedia content in determining the virality of social media information. *Information*, 3(3), 278–289. <https://doi.org/10.3390/info3030278>
- Bucher, T., & Helmond, A. (2018). The affordances of social media platforms. In *The SAGE handbook of social media*. Sage. <https://doi.org/10.4135/9781473984066>
- Bürkner, P.-C. (2017). brms: An R package for bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1). <https://doi.org/10.18637/jss.v080.i01>
- De León, E., & Trilling, D. (2021). A sadness bias in political news sharing? The role of discrete emotions in the engagement and dissemination of political news on Facebook. *Social Media + Society*, 7(4), 205630512110597. <https://doi.org/10.1177/20563051211059710>
- de Oliveira, D. V. B., & Albuquerque, U. P. (2021). Cultural evolution and digital media: Diffusion of fake news about COVID-19 on Twitter. *SN Computer Science*, 2(6), 430. <https://doi.org/10.1007/s42979-021-00836-w>
- Derczynski, L., Albert-Lindqvist, T. O., Bendsen, M. V., Inie, N., Pedersen, V. D., & Pedersen, J. E. (2019, October 31). Misinformation on Twitter during the Danish national election: A case study. *Proceedings of the Conference for Truth and Trust Online 2019*. <https://doi.org/10.36370/tto.2019.16>
- DR Analyse. (2024). *Medieudviklingen 2023*. Danmarks Radio (DR). <https://tinyurl.com/34zwmy44>
- Dunbar, R. I. M. (2004). Gossip in evolutionary perspective. *Review of General Psychology*, 8(2), 100–110. <https://doi.org/10.1037/1089-2680.8.2.100>
- Dunbar, R. I. M. (2009). The social brain hypothesis and its implications for social evolution. *Annals of Human Biology*, 36(5), 562–572. <https://doi.org/10.1080/03014460902960289>
- Ferrara, E., & Yang, Z. (2015). Quantifying the effect of sentiment on information diffusion in social media. *PeerJ Computer Science*, 1, e26. <https://doi.org/10.7717/peerj-cs.26>
- Fine, J. A., & Hunt, M. F. (2023). Negativity and elite message diffusion on social media. *Political Behavior*, 45(3), 955–973. <https://doi.org/10.1007/s11109-021-09740-8>
- Funke, D. (2020). *PolitiFact | fact-checking 'Plandemic': A documentary full of false conspiracy theories about the coronavirus*. <https://www.politifact.com/article/2020/may/08/fact-checking-plandemic-documentary-full-false-con/>
- Gabry, J., & Mahr, T. (2022). *bayesplot: Plotting for bayesian models*. <https://mc-stan.org/bayesplot/>
- Gelman, A., Goodrich, B., Gabry, J., & Vehtari, A. (2019). R-squared for bayesian regression models. *The American Statistician*, 73(3), 307–309. <https://doi.org/10.1080/00031305.2018.1549100>
- Goodrich, K. (2011). Anarchy of effects? Exploring attention to online advertising and multiple outcomes. *Psychology & Marketing*, 28(4), 417–440. <https://doi.org/10.1002/mar.20371>
- Greene, W. H. (2003). *Econometric analysis* (5th ed). Prentice Hall.
- Gross, J., & Von Wangenheim, F. (2022). Influencer marketing on Instagram: Empirical research on social media engagement with sponsored posts. *Journal of Interactive Advertising*, 22(3), 289–310. <https://doi.org/10.1080/15252019.2022.2123724>
- Hartig, F. (2024). *Installing, loading and citing the package* [Computer software]. <https://cran.r-project.org/web/packages/DHARMa/vignettes/DHARMa.html>
- Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, 53(1), 575–604. <https://doi.org/10.1146/annurev.psych.53.100901.135109>
- Hilbe, J. M. (2011). *Negative binomial regression* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511973420>
- Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018). Social learning strategies: Bridge-building between fields. *Trends in Cognitive Sciences*, 22(7), 651–665. <https://doi.org/10.1016/j.tics.2018.04.003>

- King, K. K., & Wang, B. (2023). Diffusion of real versus misinformation during a crisis event: A big data-driven approach. *International Journal of Information Management*, 71, 102390. <https://doi.org/10.1016/j.ijinfomgt.2021.102390>
- Li, Y., & Xie, Y. (2020). Is a picture worth a thousand words? An empirical study of image content and social media engagement – Yiyi Li, Ying Xie, 2020. *Journal of Marketing Research*. https://journals.sagepub.com/doi/full/10.1177/0022243719881113?utm_source=chatgpt.com
- López-García, X., Costa-Sánchez, C., & Vizoso, Á. (2021). Journalistic fact-checking of information in pandemic: Stakeholders, hoaxes, and strategies to fight disinformation during the Covid-19 crisis in Spain. *International Journal of Environmental Research and Public Health*, 18(3), 1227. <https://doi.org/10.3390/ijerph18031227>
- Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, 6(60), 3139. <https://doi.org/10.21105/joss.03139>
- Luo, H., Meng, X., Zhao, Y., & Cai, M. (2023). Exploring the impact of sentiment on multi-dimensional information dissemination using COVID-19 data in China. *Computers in Human Behavior*, 144, 107733. <https://doi.org/10.1016/j.chb.2023.107733>
- Lyons, A., & Kashima, Y. (2001). The reproduction of culture: Communication processes tend to maintain cultural stereotypes. *Social Cognition*, 19(3), 372–394. <https://doi.org/10.1521/soco.19.3.372.21470>
- Martin, D., Cunningham, S. J., Hutchison, J., Slessor, G., & Smith, K. (2017). How societal stereotypes might form and evolve via cumulative cultural evolution. *Social and Personality Psychology Compass*, 11(9), e12338. <https://doi.org/10.1111/spc3.12338>
- Marwick, A. E., & boyd, d. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, 13(1), 114–133. <https://doi.org/10.1177/1461444810365313>
- McGuigan, N., & Cubillo, M. (2013). Information transmission in young children: When social information is more important than nonsocial information. *The Journal of Genetic Psychology*, 174(6), 605–619. <https://doi.org/10.1080/00221325.2012.749833>
- Mesoudi, A., Whiten, A., & Dunbar, R. (2006). A bias for social information in human cultural transmission. *British Journal of Psychology*, 97(3), 405–423. <https://doi.org/10.1348/000712605X85871>
- Metzler, H., & Garcia, D. (2022). *Social drivers and algorithmic mechanisms on digital media* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/cxa9u>
- Metzler, H., & Garcia, D. (2023). Social drivers and algorithmic mechanisms on digital media. *Perspectives on Psychological Science*, 19(5), 735–748. <https://doi.org/10.1177/17456916231185057>
- Morin, O., & Acerbi, A. (2017). Birth of the cool: A two-centuries decline in emotional expression in Anglophone fiction. *Cognition and Emotion*, 31(8), 1663–1675. <https://doi.org/10.1080/02699931.2016.1260528>
- Mousavi, M., Davulcu, H., Ahmadi, M., Axelrod, R., Davis, R., & Atran, S. (2022). Effective messaging on social media: What makes online content go viral? *Proceedings of the ACM Web Conference 2022*, 2957–2966. <https://doi.org/10.1145/3485447.3512016>
- Nairne, J. S., Thompson, S. R., & Pandeirada, J. N. S. (2007). Adaptive memory: Survival processing enhances retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(2), 263–273. <https://doi.org/10.1037/0278-7393.33.2.263>
- Newman, N., Fletcher, R., Robertson, C. T., Ross Arguedas, A., & Nielsen, R. K. (2024). *Reuters Institute digital news report 2024*. Reuters Institute for the Study of Journalism, University of Oxford. <https://doi.org/10.60625/RISJ-VY6N-4V57>
- Nissen, I. A., Walter, J. G., Charquero-Ballester, M., & Bechmann, A. (2022). Digital infrastructures of COVID-19 misinformation: A new conceptual and analytical perspective on fact-checking. *Digital Journalism*, 10(5), 738–760. <https://doi.org/10.1080/21670811.2022.2026795>
- OSF. (2023, October 16). *Sharing, commenting and reacting to Danish Misinformation: A Case Study of Cognitive Attraction on Facebook* [Project]. https://osf.io/dqrky/?view_only=6680bf5c343344f48ead45fe7314af0e
- Pieters, R., & Wedel, M. (2004). Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing*, 68(2), 36–50. <https://doi.org/10.1509/jmkg.68.2.36.27794>
- R Core Team. (2023). *R: A language and environment for statistical computing*. <https://www.R-project.org/>

- Rathje, S., Robertson, C., Brady, W. J., & Van Bavel, J. J. (2024). People think that social media platforms do (but should not) amplify divisive content. *Perspectives on Psychological Science*, 19(5), 781–795. <https://doi.org/10.1177/17456916231190392>
- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, 5(4), 296–320. https://doi.org/10.1207/S15327957PSPR0504_2
- Scheffer, M., van de Leemput, I., Weinans, E., & Bollen, J. (2021). The rise and fall of rationality in language. *Proceedings of the National Academy of Sciences*, 118(51), e2107848118. <https://doi.org/10.1073/pnas.2107848118>
- Schöne, J. P., Parkinson, B., & Goldenberg, A. (2021). Negativity spreads more than positivity on Twitter after both positive and negative political situations. *Affective Science*, 2(4), 379–390. <https://doi.org/10.1007/s42761-021-00057-7>
- Song, X., Petrak, J., Jiang, Y., Singh, I., Maynard, D., & Bontcheva, K. (2021). Classification aware neural topic model and its application on a new COVID-19 disinformation corpus. *PLOS ONE*, 16(2), e0247086. <https://doi.org/10.1371/journal.pone.0247086>
- Statista. (2023). *Most popular social networks worldwide as of January 2023, ranked by number of monthly active users*. <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>
- Statista. (2024). *Daily social media usage worldwide*. <https://www.statista.com/statistics/433871/daily-social-media-usage-worldwide/>
- Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—Sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, 29(4), 217–248. <https://doi.org/10.2753/MIS0742-1222290408>
- Stubbersfield, J. M. (2022). Content biases in three phases of cultural transmission: A review. *Culture and Evolution*, 19(1), 41–60. <https://doi.org/10.1556/2055.2022.00024>
- Stubbersfield, J. M. (2025). Content-based learning biases. In T. Shackelford (Ed.), *Encyclopedia of religious psychology and behavior* (pp. 1–16). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-38971-9_134-1
- Stubbersfield, J. M., Flynn, E. G., & Tehrani, J. J. (2017). Cognitive evolution and the transmission of popular narratives: A literature review and application to urban legends. *Evolutionary Studies in Imaginative Culture*, 1(1), 121–136. <https://doi.org/10.26613/esic.1.1.20>
- Stubbersfield, J. M., Tehrani, J. J., & Flynn, E. G. (2015). Serial killers, spiders and cybersex: Social and survival information bias in the transmission of urban legends. *British Journal of Psychology*, 106(2), 288–307. <https://doi.org/10.1111/bjop.12073>
- TjekDet. (2020, August 4). *Arbejdsproces og etisk regelsæt [Work process and ethical code]*. Tjekdet. <https://www.tjekdet.dk/arbejdsproces-og-etisk-regelsaet>
- Tsugawa, S., & Ohsaki, H. (2015). Negative messages spread rapidly and widely on social media. *Proceedings of the 2015 ACM on Conference on Online Social Networks*, 151–160. <https://doi.org/10.1145/2817946.2817962>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Warrens, M. J. (2015). Five Ways to Look at Cohen’s Kappa. *Journal of Psychology & Psychotherapy*, 5(4), 1000197. <https://doi.org/10.4172/2161-0487.1000197>
- WHO. (2020). *Infodemic*. <http://www.who.int/westernpacific/health-topics/infodemic>
- Youngblood, M., Stubbersfield, J. M., Morin, O., Glassman, R., & Acerbi, A. (2023). Negativity bias in the spread of voter fraud conspiracy theory tweets during the 2020 US election. *Humanities and Social Sciences Communications*, 10(1), 573. <https://doi.org/10.1057/s41599-023-02106-x>