Interaction as Power and Responsibility

How Social Media Reshapes User Roles and Communication Patterns

Maximilian Haug
University of Bamberg
Neu-Ulm University of Applied Sciences

Interaction as Power and Responsibility

How Social Media Reshapes User Roles and Communication Patterns

Diese Arbeit hat der Fakultät Wirtschaftsinformatik und Angewandte Informatik der Otto-Friedrich-Universität Bamberg als Dissertation vorgelegen.

Erstgutachter: Prof. Dr. Tim Weitzel

Zweitgutachter: Prof. Dr. Heiko Gewald

Mitglied der Promotionskommission: Prof. Dr. Oliver Posegga

Tag der Disputation: 31.10.2025

URN: urn:nbn:de:bvb:473-irb-112483x
DOI: https://doi.org/10.20378/irb-112483

Table of Contents

Chapter III – Ramifications

Appendix

Acknowledgements

My journey as a PhD student, which culminated in this dissertation, was exciting and not always easy. However, I always live by the philosophy that we cannot acknowledge the truly wonderful moments if there were no hardships that we overcame to reach that point. I am very grateful to the many people who stood by me through the hardships and then celebrated the winnings.

I would like to first thank Prof. Dr. Heiko Gewald and Prof. Dr. Tim Weitzel for making it possible for me to embark on this academic journey. Doing a PhD was not really on my radar when I started my Master's program, until Heiko Gewald sparked my curiosity about information systems research. He later created a wonderful environment where I was free to explore science the way I wanted. Tim Weitzel also encouraged my research and consistently provided valuable guidance on how to focus my work, since it was very easy to go off track. Lastly, I want to thank Oliver Posegga for serving on my PhD committee.

Furthermore, I want to thank all my friends and colleagues who were and still are at the Neu-Ulm University of Applied Sciences. First, I would like to thank Dr. Robert Rockmann, my first office and research colleague, who introduced me to all things HNU from a PhD student perspective. I further want to thank Dr. Thomas Schulz, Prof. Dr. Anna Wiedemann, and Prof. Dr. Andy Weeger for providing valuable insights in my first years. Huge thanks go to Nikola Finze-Pfaff, Luca Stegmann, Deinera Jechle, Kathrin Kim, Dr. Axel Hund, and Dr. Tanja Schröder, who actively pushed me to further heights in research and beyond. I could not imagine better colleagues. Also, thanks to all the wonderful people who started their journey at HNU when I was about to leave, namely, Dr. Adeline Frenzel-Piasentin, Markus Bink, Luisa-Sophie Gloger, and Nils Riekers.

Special thanks go to Prof. Dr. Marten Risius, Prof. Dr. Christian Maier, Prof. Dr. Jason B. Thatcher, and Prof. Dr. Heinz-Theo Wagner, who have always provided feedback on my papers when I asked them. Your feedback was always critical and honest.

Lastly, I owe so much to my family and friends beyond academia. First of all, thanks to my parents, Maria and Werner. I do not know what else a child could wish for. You were always supportive of what I did. You gave me the security and love for me to succeed. I also want to thank my brother Raimund, who is one of my best friends and with whom I share so many memories. Importantly, I want to thank my girlfriend, Laura, for supporting me through my PhD journey. You were curious about how you helped me, and I was just confused about how you did not see how much you helped me. The same applies to my dearest friends: Maren, Lara, Alissa, and Cynthia. Without the joy you brought to my life, this would not have been possible. Please stay as you are and let us experience even more fun.

Thank you all.

Zusammenfassung
(German Summary)

Soziale Medien haben den Informationsaustausch in unserer Gesellschaft grundlegend verändert. Nutzer interagieren täglich weltweit mit anderen Nutzern, ohne an physische oder zeitliche Barrieren gebunden zu sein (Karahanna et al. 2018). Das Aufkommen sozialer Medienplattformen hat die Kommunikation in zweierlei Hinsicht erschüttert. Erstens fand ein Strukturwandel in der Kommunikation statt. Soziale Medien führten eine Multiplexität in der Kommunikation ein, die den unmittelbaren Zugang und die Interaktion mit anderen Menschen kennzeichnet. Diese Form des Zugangs unterscheidet sich erheblich von der traditionellen Kommunikation. Soziale Medien ermöglichen beinahe uneingeschränkte Viele-zu-Viele-Verbindungen (Larson and Watson 2011; Treem and Leonardi 2012). Unterschiedliche Plattformen bieten beispielsweise Funktionen wie das „Folgen" anderer Nutzer oder das Teilen von Inhalten, das sogenannte „Sharen". Solche Funktionen ermöglichen es Nutzern, Kommentare in Echtzeit zu sehen und beliebte Nachrichten schnell zu verbreiten, was zu einer schnellen Informationsverbreitung führt (Venkatesan et al. 2021).

Die Stärkung der Nutzer, und der damit eingehende Rollenwandel ist zentral für das Verständnis der modernen Kommunikation. Nutzer entwickeln sich vom passiven Informationskonsumenten zum aktiven Inhaltsverteiler. Daher beschäftigt sich die vorliegende Arbeit mit der Interaktion („Like, Comment, Share") von Nutzern auf Social Media und verfolgt das übergeordnete Forschungsziel:

Forschungsziel: Wie verändern soziale Medien die Kommunikation durch Nutzerinteraktion?

Die Struktur der Arbeit baut auf einer Kaskade von Veränderungen Strukturen in der Kommunikation (Kapitel 1), hin zur Interaktion (Kapitel 2), und letztlich deren Auswirkungen auf den individuellen Nutzer (Kapitel 3).

Das erste Kapitel der Arbeit fokussiert sich hierbei auf die strukturellen Veränderungen, die soziale Medien in der Kommunikation hervorgerufen haben. Hierbei werden insbesondere die sich veränderten Rollen der Nutzer, Plattformen, und traditionellen Nachrichten, sowie Meinungsführer eingegangen. Die Ergebnisse des ersten Kapitels zeigen, dass soziale Medien ihren Nutzern Handlungsspielraum bieten, aber auch erhöhte Verantwortung mit sich bringen. Nutzer haben ein hohes Potenzial, das Online-Informationsökosystem zu ihren Gunsten zu gestalten. Die Ergebnisse zeigen jedoch, dass Nutzer ihren Handlungsspielraum nicht ausreichend berücksichtigen, da ihr Verhalten hauptsächlich auf Heuristiken beruht, die auf der Wahrnehmung von Profilen basieren, die Inhalte teilen, an denen sie interessiert sind. Daher müssen sich Social-Media-Nutzer ihrer Verantwortung bewusstwerden, um ihren neu gewonnenen Handlungsspielraum effektiv zu nutzen.

Das zweite Kapitel beschäftigt sich mit dem Hauptfokus der vorliegenden Arbeit: Nutzerinteraktion. Hierbei wird auf Interaktion auf Social Media als Ganzes eingegangen, sowie auf Charakteristiken von Interaktion in Social Media Gruppen (z.B. Facebook Groups, SubReddits). Die Ergebnisse zeigen, wie

die identifizieren drei Haupttypen von Accounts auf Social Media Interaktion beeinflussen. Die drei Typen sind: Nachrichtenmedien, Meinungsführer, und normale Nutzer. Hierbei zeigt sich, dass durch das Teilen der Inhalte und dem Fokus auf Heuristiken ein sogenannter „Messenger Effekt“ ausbildet. Der Account, der Inhalte teilt (und nicht der ursprüngliche Urheber), beeinflusst die Wahrnehmung je nach wahrgenommener Ähnlichkeit. Zusätzlich zeigt sich ein „Mixed-Actor-Amplification Effekt“, welcher zeigt, dass sollte der Urheber zusätzlich unähnlicher sein, steigt die Glaubwürdigkeit des Inhalts noch weiter an. Weiterhin zeigen die Ergebnisse eine gewisse Hierarchie der Einflüsse, ob Nutzer mit Inhalten interagieren. Bestätigungsfehler (Comfirmation Bias) und Ähnlichkeit bestimmen zwar die Grundlage der Wahrnehmung von Inhalten, können jedoch durch sozialen Druck überlagert werden. Entsprechend identifiziert diese Arbeit unterschiedliche Grundlagen, die zu sozialem Druck führen können, wie die Angst vor Isolation, Identität innerhalb der Gruppe, und soziale Unterstützung. Letztlich zeigen die Ergebnisse dazu, dass sogenannte Klick-Interaktion (z.B. „Liking“) niedrigschwellige Handlung darstellt, die häufig genutzt wird, um Zugehörigkeit zur Community zu signalisieren – auch ohne inhaltliche Zustimmung. Zusätzlich wird die Rolle des Themas der Inhalte beleuchtet. Während Kommentieren vom Thema und seiner Wichtigkeit abhängt, ist die Klick-Interaktion vor allem sozial motiviert.

Im dritten Kapitel beleuchtet die Arbeit die negativen Seiten Auswirkungen von Social Media Interaktion. Hierbei liegt der Fokus auf Interaktion, welche Stress in Nutzern auslösen kann. Zwei Hauptursachen werden identifiziert: Erstens entsteht Stress in sozialen Gruppen durch soziale Überlastung, wenn Mitglieder aktiv Unterstützung leisten. Es zeigt sich, dass nicht die Erwartung, sondern das tatsächliche Unterstützen belastend wirkt. Zweitens können Interaktionsmetriken wie Likes und Kommentare Stress verursachen. Nutzer erwarten eine bestimmte Anzahl an Likes, um eine Interaktion als positiv zu werten. Wird diese Schwelle nicht erreicht, empfinden sie die Resonanz als negativ. Likes sind dabei zentraler Bewertungsmaßstab, während fehlende Kommentare weniger bedeutsam sind. Zur Stressbewältigung nutzen Nutzer verschiedene Strategien wie emotions- und problemorientierte Coping-Ansätze, sowohl online als auch offline. Negativ wahrgenommene Interaktionen lassen Emotionen wie Wut und Traurigkeit entstehen. Diese spiegeln das Missverhältnis zwischen Erwartung und Realität wieder. Während wutbasierte Strategien kaum entlastend wirken, fördern traurigkeitsbasierte Ansätze häufig weiteres Engagement – etwa durch neue Beiträge oder Interaktionen. Die Arbeit zeigt damit, dass selbst negative Emotionen zu verstärkter Social Media-Nutzung führen können.

Zusammenfassend lässt sich feststellen, dass der normale Nutzer auf Social Media eine bedeutendere Rolle in der modernen Kommunikation zufällt. Als Aktiver Verteiler von Inhalten durch Interaktion müssen sich Nutzer ihrer neuen Freiheiten und Verantwortungen stellen, um für sich, wie auch für die Gesellschaft eine nachhaltige und belastbare Informationsumgebung zu schaffen.

Introductory Paper

Interaction as Power and Responsibility

How Social Media Reshapes User Roles and Communication Patterns

Maximilian Haug
University of Bamberg
Neu-Ulm University of Applied Sciences

Interaction as Power and Responsibility

How Social Media Reshapes User Roles and Communication Patterns

1 Introduction

Social media substantially changed how our society exchanges information. Users interact with other users on a daily basis without being bound to physical or time barriers all over the world (Karahanna et al. 2018). User empowerment is central to understanding how communication has evolved. This empowerment can be found in the shifting role of users from mere passive information consumers to active content distributors. This dissertation argues that user interaction on social media is not merely a behavioral artifact but a structurally embedded phenomenon that redefines communication roles and outcomes.

The emergence of social media platforms shocked communication in two different ways. First, we perceived a structural change in communication. Social media introduced a multiplexity in communication with the characteristics of immediate access and interaction with other people. This form of access is vastly different from traditional communication, in which technology did not provide the possibility to reach out to millions of individuals. Social media enables many-to-many connections (Larson and Watson 2011; Treem and Leonardi 2012) without restrictions. Twitter (now X), for example, includes functionalities such as "following" other users or sharing tweets known as "retweeting." Such functionalities enable Twitter users to see comments in real-time and enable the broadcasting of popular messages quickly, which results in rapid information dissemination (Venkatesan et al. 2021).

Second, we observe a societal change in the form of newly emerging and empowered roles on social media platforms. Traditionally, gatekeepers (e.g., news organizations) that we associated with specific journalistic standards acted as a filter of information for the masses (Westley and MacLean Jr 1957). This was due to their unique characteristic of solely having access to a big target audience. Now, with the rise of social media platforms, we perceive that gatekeepers are increasingly de-emphasized or discredited. On the other side, we see the emergence of new roles, such as opinion leaders (e.g., politicians, influencers) with a clear agenda (Sahelices-Pinto and Rodríguez-Santos 2014). As a result, there is potential in social media for a narrowing of information diversity and a partisan shift in news consumption (Kitchens et al. 2020). Since opinion leaders can now directly communicate with their target audience, gatekeepers have lost one of their key characteristics, access to the masses, in information dissemination. Lastly, and arguably one of the most important changes, the audience, now ordinary social media users, has been greatly empowered by the communication process. They are now an active part of the information dissemination process through interaction. Interaction in this case refers to social media inherent mechanisms such as “liking”, “commenting”, or “sharing”. Liking, for

example, is a statement of endorsement, but also a mechanism that further disseminates content as the network of the user also sees that content was liked. Commenting is a behavior that actively creates new content on social media with which other users can interact. Hence, content refers to types of information posted on social media (e.g., text, video, audio, imagery, etc.) by their users. Sharing actively spreads content to the personal network. Hence, users are no longer passive consumers of content, but when they start interacting with content, they are the ones who further disseminate it.

These changes have vast ramifications for communication in our society. One of the most prominent is disseminating false information on social media, which is seen as one of the largest global concerns (Wei et al. 2019) that is elevated by users further spreading false information or narratives (Starbird et al. 2023). Recent research showed that disinformation campaigns are largely based on authentic user interaction in opposition to bot networks (Starbird et al. 2023). Furthermore, opinion leaders disseminate information without the need to adhere to integrity or accuracy, resulting in disinformation campaigns (Linvill et al. 2019; Linvill and Warren 2020) as users become such a crucial part in the information dissemination process through their interaction. Information Systems (IS) research proposes traditional theories, such as Media Synchronicity Theory, which emphasize synchronous communication but were originally developed for linear, one-to-many media environments (Dennis et al. 2008; Shannon 2001; Westley and MacLean Jr 1957). Research has often applied these linear models to social media contexts (e.g., George et al. 2018), failing to fully account for the emerging interaction patterns between users (Bennett and Segerberg 2012; Citron 2019; Milan 2015; Vaast et al. 2017). This dissertation stresses the change from passive consumption by the users within the audience towards active participation. Hence, the overarching research question for this dissertation is:

How does social media change communication through user interaction?

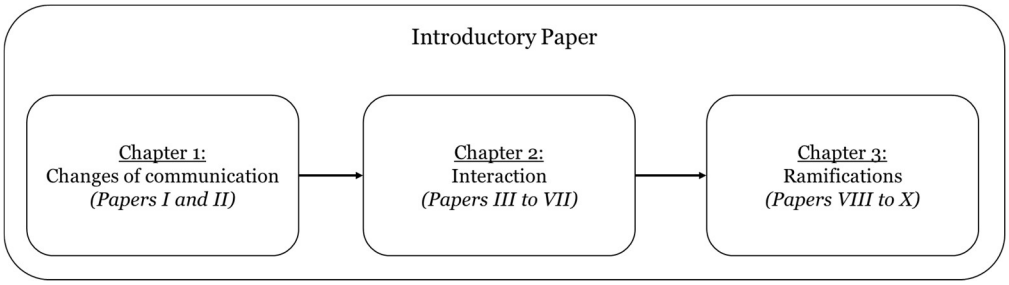


Figure 1. Structure of the dissertation with an introductory paper and three main chapters

To answer this research question, this dissertation undertakes a three-step approach to understand origins, the resulting interaction (social media as a whole and within social media communities), and the ramifications of interaction. Therefore, this dissertation provides a comprehensive cascade from changes within communication to the outcomes stemming from these changes.

This dissertation provides an introductory paper that consolidates ten papers that tap into different aspects of this cascade. It provides the theoretical foundations, the methodology used, individual findings, and aggregated contributions. The introductory paper further provides a literature review with a focus on information systems (IS) research to derive the research gaps. The following ten papers are

then structured into three different chapters. Each paper provides insights beyond its chapter, but mainly focuses on providing answers to its chapter. The first chapter focuses on the changes in communication introduced by social media platforms (**Papers I** and **II**). The second chapter focuses on the interaction stemming from the new role of the user as the audience in different contexts. This chapter has two sub-chapters. Sub-chapter one observes interaction on social media as a whole (**Papers III** to **V**), while sub-chapter two has the explicit perspective of social media communities (**Papers VI** and **VII**). Chapter four finally highlights the ramifications stemming from such interaction and focuses on unintended consequences on the level of the user (**Papers VIII** to **X**).

2 Theoretical Foundations

This section contains background knowledge about communication processes and theoretical concepts to understand the results of the individual papers.

2.1 The Change of Communication in the Era of Social Media

The communication process generally involves a sender and a receiver. Person A transmits something to person B about an object X (Newcomb 1953). This relatively simple concept can become very complex if multiple actors are involved in the dissemination of information. In our current communication environment, which is dominated by communication via the internet, social media introduces various actors and roles that shape and alter communication from what we traditionally know. To understand the changes within the online information ecosystem, it is first crucial to understand traditional communication models, such as the model of Westley and MacLean Jr (1957). The initial goal was to untangle a "jungle of unrelated concepts and systems of concepts" (Westley and MacLean Jr 1957, p. 8) in the various research streams of mass communication. Their model refers to different research results on communication roles and patterns. The model incorporates senders and receivers, along with gatekeepers who initially had the only means to disseminate information to a big number of recipients. Thereby, gatekeepers had a powerful role in what the masses were able to read (e.g., newspapers), hear (e.g., radio), or watch (e.g., TV). Even though the model has received valid criticism, such as the passive character of the audience (e.g.,Van Dijck 2009), as it characterizes the audience as passive recipients, whose levels of passivity may vary (Bolin 2012), the model still shows valid insights. The model provides a useful framework for researchers to identify the mechanisms of mass communication.

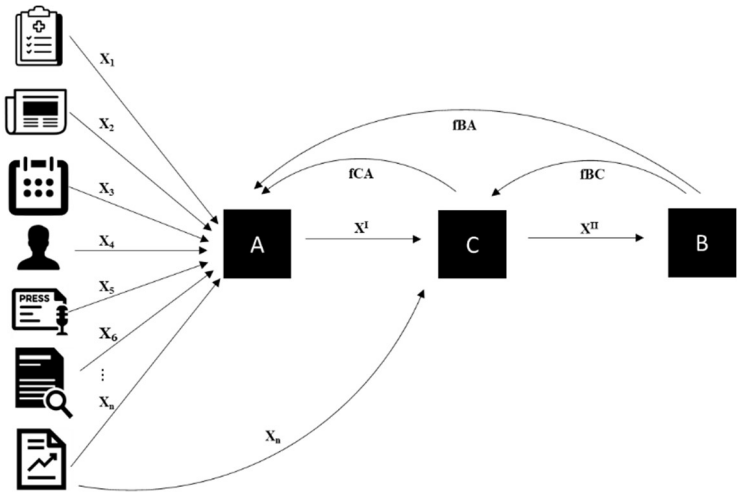


Figure 2. Communication model by Westley and MacLean Jr (1957)

Westley and MacLean Jr (1957) described communication as a process where information is filtered before reaching its audience. In their model, actors A or C select details about events or sources to communicate through organized formats such as press releases (X1-Xn). Actor A, who can be any type

of communicator (e.g., politicians), sends information (XI) to C (e.g., news outlets), who then delivers it to the audience (B) using mass media channels (XII) like TV, radio, or newspapers (McQuail and Windahl 2015). The news media play the role of the gatekeeper, as they filter and reshape the information before its dissemination to the public in a one-to-many communication process. While media organizations like C have a broad reach, the audience is typically limited to personal or interpersonal communication. The model also incorporates feedback loops between B and C (fBC), B and A (fBA), and C and A (fCA) throughout the communication process.

To understand how information spreads within online information ecosystems, Westley and MacLean Jr (1957) provide a solid basis that can be adapted to the context of social media. Although the basic communication process (A -> C -> B) remains unchanged, social media's inherent characteristics have significantly changed the communication landscape. Here, the literature within journalism and mass communication points out the influence of technology in shaping communication processes. The technological influences here refer especially to algorithms and nonhuman actors, such as bots and automated systems (Guzman and Lewis 2020; Lewis and Westlund 2015).

In recent years, social media platforms have transformed communication ecosystems by offering vastly unregulated opportunities for interaction, self-expression, content sharing, and communication (Karahanna et al. 2018). Unlike traditional face-to-face communication, social media now allows audience members to distribute content more widely (Bansal et al. 2020; Sun et al. 2022). They also have greater autonomy in verifying the quality of news compared to participants in offline communities (Xiao 2022). Social media enables various actors — individuals or organizations creating, sharing, and distributing content — to engage with large audiences (Maweu 2019). As a result, audiences have shifted from being passive consumers, as seen in traditional mass communication (e.g., one-way communication like TV and radio), to becoming active creators of user-generated content (Van Dijck 2009), to the point that they become influencers with big audiences. Furthermore, while traditional mass communication research treated technology as lacking agency, recent studies recognize technology's active role in communication processes, as algorithms and content management systems increasingly influence decision-making and vastly determine the information diet of the platform users (Guzman and Lewis 2020). Hence, we see a shift in roles and communication patterns introduced by social media.

2.2 Spaces and Types of Interaction on Social Media

As mentioned, users within the audience are now empowered to interact with each other on a grand scale and effectively act as content creators and disseminators due to their interaction based on content from other users. Social media provides opportunities for its users to interact with content on different levels. There is the possibility for content sharing and communication (Karahanna et al. 2018; Kietzmann et al. 2011). Such interaction with content and other users can happen in spaces that are considered "open" or "closed". The open context consists of the publicly available content that is posted on social media. In the closed context, the possibility of social media to create groups or communities is referenced (Karahanna et al. 2018; Kietzmann et al. 2011). These communities are distinct spaces on social media platforms that can be open to join for others, closed (i.e., approval to join is required), or

secret (i.e., an invitation is needed) (Kietzmann et al. 2011). Such communities naturally have their own rules and topics that users discuss. Therefore, we can distinguish between interaction happening in the open space of social media and within distinct communities.

Within open or closed spaces, users can then interact with others' content based on some basic mechanisms introduced by social media platforms. First, we find the most basic function, commenting, that was also prevalent in forum spaces before the emergence of social media (Gabbiadini et al. 2013; Karahanna et al. 2018). Commenting refers to interaction based on actual opinion disclosure in the form of text. However, research shows that only a small portion of users online actively take part in such interaction (Gabbiadini et al. 2013).

A distinct new mechanism of interaction, and what primarily drives social media interaction with content, is considered to be so-called click speech (Robbins 2014; Sklan 2013). Click speech via liking or sharing provides an effortless mechanism to interact with content simply via one click. Furthermore, click speech is a salient metric on which users can evaluate how popular presented content is on social media (Van Aelst et al. 2017). Sharing proposes a mixture of click speech and commenting, since it is possible to simply share content without further consideration, or to share content while providing further commenting (Pang et al. 2016; Sklan 2013). The mechanisms of liking and sharing can be found on the major social media platforms, but can vary in terms of how they are implemented. For example, sharing on X (formerly Twitter) would be considered re-tweeting, and liking in the context of Reddit would be a so-called “upvote”.

2.3 Main Drivers of Interaction on Social Media

2.3.1 Confirmation Bias

When individuals interact with content on social media, they are often led by confirmation bias. Humans are cognitive misers and let themselves drive in terms of social engagement through previously held beliefs (Kim and Dennis 2019). This means information that confirms their beliefs is automatically deemed as believable or more valuable, while ignoring information that challenges them (Devine et al. 1990; Koriat et al. 1980). This effect generally emerges together with another cognitive bias, namely, cognitive dissonance (Festinger 1962). In the case of information that does not reaffirm previously held beliefs, individuals need to resolve this form of cognitive dissonance. Cognitive dissonance is the mental unease or tension evoked by confronting challenging information. To get rid of the cognitive dissonance, individuals reject contradictory information and overemphasize reaffirming information (Simon 1979).

Studies have connected confirmation bias to the dismissal of counterinformation, especially in research concerning false information (also in fake news research). Confirmation bias can overwrite credible sources (Halbach et al. 2019; Moravec et al. 2019) and lead to the refusal to call out false information (Coscia and Rossi 2020). Furthermore, flagging mechanisms on social media show lower efficacy when individuals engage with posts that do not align with their beliefs (Moravec et al. 2019). Contrary to that, if individuals engage with content that reaffirms their beliefs, they feel more confident in their beliefs, and the perceived truthfulness of the information is elevated (Halbach et al. 2019).

2.3.2 Privacy

In literature, privacy is commonly viewed as the ability of individuals to regulate the disclosure of their personal information to others. Assessing the impact of privacy remains a challenge due to the inherent difficulty in quantifying privacy itself. As a result, researchers often rely on related indicators like awareness, concerns, perceived risk, and trust to evaluate how privacy affects people's actions and decisions (Malhotra et al. 2004; Smith et al. 2011). These influences seem to change based on the context since risks can be perceived differently. Social media now provides a unique communication environment in which individuals can disclose personal information. Due to the nature of such information, this disclosure can be subject to privacy issues. For example, disclosing personal health status while discussing health-related issues with peers over social media.

The literature indicates that privacy-related concepts are strongly dependent on how individuals perceive them. This conclusion can be inferred by the existence of two privacy perspectives on how privacy ultimately influences behavior: the privacy calculus and the privacy paradox.

The privacy calculus perspective posits that individuals engage in an internal cost-benefit analysis. They weigh the advantages of a behavior against the associated privacy risks or costs (Klopfer and Rubenstein 1977; Posner 1981). A main pillar of the privacy calculus is the perception of risks (Dinev et al. 2006). The privacy calculus faces criticism, since individuals still engage in behaviors despite low benefits and high privacy risks (Acquisti 2004). Furthermore, literature describes a sense of resignation among users, who acknowledge the threat but feel powerless to control it. This resignation can distort the perceived balance between risks and benefits (Wirth et al. 2018).

The privacy paradox describes the notion that individuals engage in behavior that they deem problematic in terms of privacy, but continue such behaviors (Norberg et al. 2007). For example, they may perceive social media platforms as selling their personal data, but remain on the platform. This paradox has been extensively studied in the context of social media platforms. Acquisti and Grossklags (2005) provide an explanation for the paradox since individuals may discount long-term privacy risks or simply underestimate them. In return, they value the immediate benefits much more. Smith et al. (2011) support this idea and also provide evidence that people give more weight to present benefits than future risks. Gross and Acquisti (2005) investigated Facebook users' privacy settings. Even though participants articulated potential privacy risks, they did not change the default privacy settings on the platform. Alashoor and Baskerville (2015) introduce deep involvement as an explanation for why users perceive benefits to a greater degree than privacy concerns. They do so by drawing on the concept of cognitive absorption and arguing that users become immersed in their use, so privacy risks diminish.

The last example of paradoxical behavior arises within service or technology personalization. Privacy concerns create an issue for companies, since users with the highest demands for transparency are often the least willing to be profiled. Therefore, it becomes increasingly difficult to provide personalized services (Awad and Krishnan 2006). In summary, privacy research has consistently identified paradoxical user behavior, with significant implications for both individuals and firms.

2.3.3 Social Influence

One main determinant of opinion disclosure can be found within social influence. In the literature, various concepts can be summarized under social influence, such as herd behavior (Sun 2013) or subjective norms (Venkatesh et al. 2003). Herd behavior is influenced by perceiving a peer group that has already adopted some behavior and has been shown to be beneficial (Sun 2013). Therefore, a larger group of peers, that is rather faceless, determines behavior. Subjective norms refer to the perception of people who are important to oneself and their behavior, which ultimately will be adapted (Venkatesh et al. 2003). Therefore, for herd behavior, the larger the group, the stronger the effect, while for subjective norm, the stronger the tie to important individuals, the stronger the effect.

Literature on fake news explicitly shows the circumstances under which individuals share information. Since social media platforms allow users to maintain relationships, users use them heavily (Karahanna et al. 2018). In practice, such relationships can be maintained by sharing content and endorsement mechanisms such as liking (Ferm and Thaichon 2023). This dissertation also digs into the importance of such click speech mechanisms and that users try to maximize them. Lastly, status seeking among peers has been shown to strongly influence sharing (Park and Lee 2020).

In summary, as the term social media suggests, information disclosure is heavily socially driven. Various concepts have already been investigated to better understand the sharing of information. This dissertation investigates further information sharing, focusing on a specific type of community peer pressure. This is important since nowadays, with social media, users can easily find distinct communities congruent with their worldview. Hence, we must assume a more substantial polarized populace, since cross-cutting content is diminished. The following introduces the spiral of silence theory with a distinct influencing mechanism: fear of isolation.

2.4 Spiral of Silence Theory

When it comes to voicing opinions, research provides different mechanisms under which opinion disclosure can happen. Within the context of social media, we find various present opinions and groups of individuals who gather under certain opinions in communities. This context of social influence and opinion disclosure is laid out within the spiral of silence theory. The spiral of silence theory is based on what individuals perceive as public or mainstream opinion on a given issue (Noelle-Neumann 1974). It suggests that individuals who recognize their opinions differ from the majority may fear social rejection and isolation, making them less likely to openly voice their views. On the other hand, those whose opinions align with the majority are free from this fear and more inclined to express their views publicly. Hence, the theory provides an explanation not only for opinion disclosure but also for self-censorship. As a result of this logic, public opinions are reinforced and disclosed, while opinions that individuals deem as not popular are silenced by the individuals themselves (Kushin et al. 2019).

Concerning the application of the theory, research showed that the theory holds in offline (Moy et al. 2001; Noelle-Neumann 1993) and online contexts (Wu et al. 2020). Since the spiral of silence relies on the perception of public opinion, the theory is inherently connected to a specific topic in question. In

that regard, research tested the theory on a range of issues such as elections (Kushin et al. 2019), foreign intervention (Neuwirth et al. 2007), and police discrimination (Fox and Holt 2018). This dissertation further increases the application of the theory to further topics.

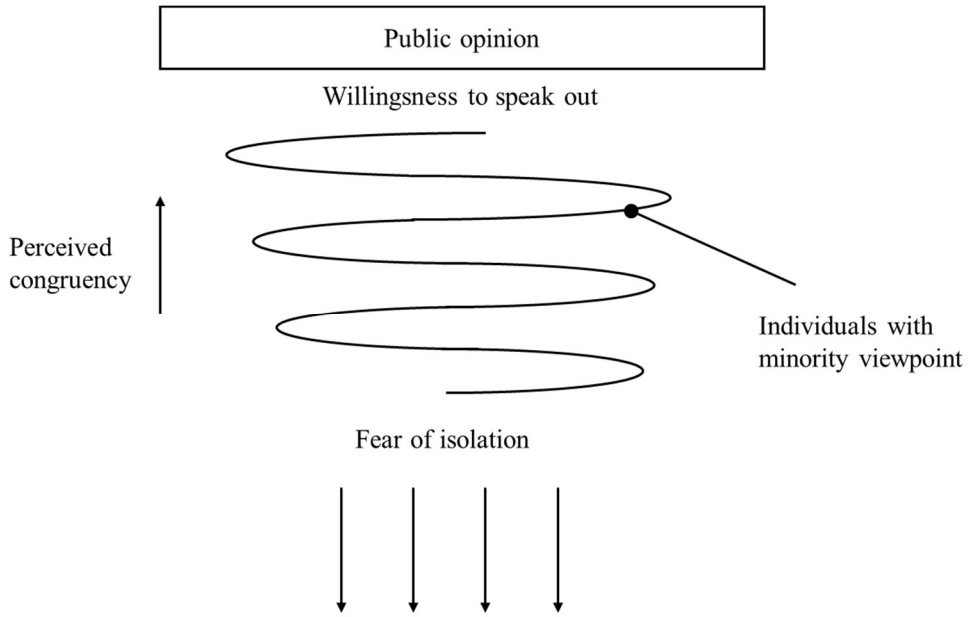


Figure 3. Spiral of silence theory illustration

A major concept of the spiral of silence is public opinion. This concept refers to the perceived opinion that compels conformity of attitude and behavior (Noelle-Neumann 1977). Individuals engage in an assessment of their social environment and determine whether their opinions on issues are consistent with public opinion using what is referred to as a "quasi-statistical organ" (Noelle-Neumann 1977). If individuals perceive their opinion to be congruent with the public opinion, they are more likely to disclose their opinion. The presence of public opinion increases their self-esteem, and they feel safe in voicing their opinion. In contrast, if their own opinion is not congruent with the public opinion, individuals feel reservations about voicing it. This reservation stems from the perception that voicing the opinion will result in social sanctions. This social sanction is described in the spiral of silence theory as social isolation from individuals who share public opinion, which is perceived as the most common opinion.

Social media provides the possibility to create distinct communities (e.g., Facebook groups, SubReddits, etc.). Within these communities, dominant community opinions can emerge, which are not necessarily congruent with what would be considered public opinion. However, such community opinions do not exist in a vacuum; they can be heavily informed by the public. For example, communities with a strong anti-establishment sentiment may take the opposite of each mainstream opinion, hence making them not detached from public opinion, while still voicing opinions not in line with public opinion.

The premise of the spiral of silence theory states that humans in nature fear separation and isolation from society and act in accordance to avoid said isolation due to humans being inherently social creatures (Noelle-Neumann 1974). To avoid such isolation, individuals carefully assess whether their disclosure of opinions might lead to a separation from society because they are unpopular or

controversial. Research states that there are two particularities of such fear of isolation. On the one hand, fear of isolation is a trait-based influence based on individual differences (Hayes et al. 2013). This means there is an inherent personality characteristic that makes individuals generally susceptible to the fear of social isolation. Personality traits are stable over time and context-independent (Scheufele et al. 2001; Thatcher et al. 2018).. Hence, trait-based fear of isolation is a constant issue for said individuals, independent from context, and therefore, they are more careful to disclose opinions that do not align with public opinion (Hayes et al. 2013). On the other hand, state-based fear of isolation is an emotional state of fear of social isolation that emerges through experiencing a certain situation (Wu et al. 2020). For example, within a conversation with an individual, a controversial topic comes up, which then triggers the state-based fear of isolation (Fox and Holt 2018; Neuwirth et al. 2007). This emotional state vanishes after the conversation and is not inherently tied to the person (Neuwirth et al. 2007). In that regard, individuals with a strong trait-based fear of isolation may experience the emotional state of fear of isolation more often (Wu and Atkin 2018). Lastly, though, both mechanisms, trait-based and state-based fear of isolation, can influence behavior.

2.5 Social Support and Social Overload

As outlined before, individuals have a fear of being isolated from communities. Therefore, it is also important to elaborate on which benefits members within online communities derive from their interaction. Social support theory outlines how social networks provide an environment in which individuals receive support to cope with negative situations. They perceive the network they are part of as supportive and therefore have a higher capacity to deal with personal problems, negative experiences, etc. (Caplan 1974; Cassel 1976; Cobb 1976). Social support is a complex and wide-ranging concept that involves various forms of assistance and encouragement provided by an individual's social circle. It has been shown that social support is vital for boosting mental and physical well-being, particularly in periods of stress or difficulty. As mentioned, social support has different facets due to its ability to deliver emotional, informational, and practical assistance, which together create a sense of security and belonging. At its heart, social support includes both the belief and the reality that one is valued. This means that individuals within the social circle have a certain perception that they have a strong attachment to the self. Furthermore, individuals can depend on others for help, and are part of a nurturing network. Therefore, there is a difference between people who are simply around and people who truly care and try to help (Cobb 1976).

As mentioned, social support can have different facets. A distinction can be made between network characteristics, perception, and action. Three types of social support are theorized in literature: embedded, perceived, and enacted social support (Caplan 1974; Cassel 1976; Cobb 1976). Embedded social support refers to how well the connection between individuals is established. Network size and density play an important role in determining embedded social support (Barrera Jr 1986; Thoits 1995). Furthermore, perceived social support only accounts for the individual's perception of the community theoretically being able to provide social support (Barrera Jr 1986; Johnsen 2001; Thoits 1995). So, no actions need to be taken, but solely the feeling of being able to rely on the network if something bad were to happen already provides value in the form of social support in itself. Finally, enacted social support

describes actual actions. Members of the social network actively provide social support by trying to talk to the individual who is in need, help out with problems, etc. (Barrera Jr 1986; Seidman et al. 2006; Thoits 1995).

The exploration of social support within social media reveals positive aspects in the interaction of users with social media communities. Social support on social media is especially relevant for communities that contain vulnerable groups. For example, in online mental health, users are embedded in a web of support through either their passive consumption of information or active information sharing in threaded discussions (Parameswaran and Kishore 2018). The interactions within social media and online mental health communities, particularly on platforms like Reddit, exhibit a rich and multifaceted structure. To better understand how different types of responses and linguistic indicators can impact users' future risk of suicidal thoughts, researchers employed stratified propensity score matching (Chen and Xu 2021). This method allowed us to analyze how distinct forms of feedback, such as emotional support (e.g., marked by empathy and expressions of concern) and informational support (i.e., which includes practical guidance and resource sharing), affect users over time.

However, social support also has a dark side. Maier et al. (2015) theorize social overload within online social networks as a downside to how users can benefit from social media communities. Social Overload is a phenomenon occurring when individuals are faced with overwhelming social demands from their peers, or when they may perceive themselves as obliged to deliver overwhelming social support. Hence, the quantity and intensity of social interactions pass an individual's capacity to actually deliver adequate social support, resulting in a draining mind state (Maier et al. 2015). The concept was first investigated in offline scenarios in which individuals were part of a dense social environment. Literature suggests that people perceive social overload when the maintenance of social relationships surpasses their individual capacity, which then infringes on the ability to interact effectively and to deliver social support (McCarthy and Saegert 1978). Now, social overload seems to be elevated within the social environment on social media. Indications are within the size that social media networks propose. Hence, user networks have grown, more social connections were established, and therefore, more requests for social support emerged (Amichai-Hamburger et al. 2013; Krasnova et al. 2010). Therefore, literature suggests that with such an increase in size and number of connections, individuals feel increasing strain from satisfying the ever-demanding social support from such social networks (Maier et al. 2015). Therefore, social media platforms can provide possibilities to receive social support, but in turn, they can also be hosts to social overload due to the increasing demands from other social media users.

2.6 Coping and Its Application within Social Media

When we talk about negative consequences arising from the use of social media, one needs to ask how individuals deal with such experiences. This question can be summarized under the term of coping with stress. In the literature, there are different theories that focus on stress and its ramifications. For example, the stress theory put forth by Selye (1978), technostress in IT contexts (Ragu-Nathan et al. 2008), the Situation-Organism-Behavior-Consequence framework (Bandura and Walters 1977), and the transactional model of stress and coping (Folkman and Lazarus 1985). Selye (1978) theory describes a universal approach that is based on the basic physical stimuli that can induce stress. Technostress

outlines stress resulting from interactions with technology (Ragu-Nathan et al. 2008). Bandura and Walters (1977) describe a Situation-Organism-Behavior-Consequence framework that can explain how individuals cope with stress. Lastly, Folkman and Lazarus (1985) propose a model that starts with stress and moves to the coping process. Stress in all the mentioned theories always comes down to a mismatch between the psychological state and the environment of the individual. Folkman and Lazarus (1985) integrate such significant aspects in their transactional model, which is why, in the following, this dissertation uses the transactional model of stress and coping as a theoretical basis.

The transactional model of stress and coping by Folkman and Lazarus (1985) describes the process from individuals experiencing stressful situations through cognitive appraisal and coping stages. Initially, individuals evaluate whether a situation causes stress and how such an experience ultimately affects their well-being. When it comes to stress, individuals normally have a certain capacity or resources on which they can rely to buffer stress. However, stressors can become overwhelming and ultimately outweigh the resources. As a result, individuals strive to find a strategy to deal with or cope with the excessive stress they experience. If stress exceeds individual resources, they seek specific coping strategies to overcome the stress. In short terms, coping describes the cognitive and behavioral measures employed to manage, reduce, master, or tolerate the internal and external demands of the person-environment transaction perceived as taxing or exceeding one's resources (Folkman et al. 1986, p. 572).

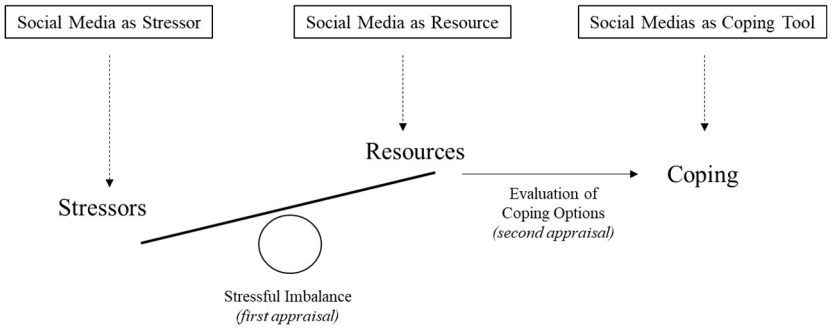


Figure 4. Transactional Model of Stress and Coping based on Wolfers and Utz (2022)

Literature proposes two approaches to how individuals cope with stress: problem-focused coping and emotion-focused coping (Folkman and Lazarus 1985). Emotion-focused coping seeks to manage emotions in response to stress in the early stages. Distraction and self-soothing are commonly employed to promote emotional relief and a sense of calm. Furthermore, problem-focused coping involves addressing the source of stress. This approach requires cognitive effort to analyze the situation, identify its underlying causes, and implement solutions that tackle the stressor directly. From this distinction, it follows that emotion-focused coping is effective for managing short-term stress, while problem-focused coping, though more cognitively demanding, provides long-term stress relief. This perspective aligns with findings by Liang et al. (2019), which suggest that emotion-focused coping often precedes problem-focused coping.

To set coping theory in context for this dissertation, there have been investigations of coping in the sphere of social media. Previous research suggests that social media can act as a coping tool (Van Ingen

et al. 2016). Social media affords the opportunity to connect with other users and exchange information about the self and beyond (Karahanna et al. 2018). As social support theory suggested beforehand, social networks can provide social support (Caplan 1974; Cassel 1976; Cobb 1976); thus, social media is also utilized as a source of social support to cope (Nabi et al. 2013). In accordance with the distinction of coping strategies, social media can be used as an emotion-focused or problem-focused coping tool (Neubaum et al. 2014; Wolfers and Utz 2022).

3 IS Literature on Social Media Interaction

Social media platforms provide audiences with the possibility to provide and engage with content (Karahanna et al. 2018). This new empowerment for the audience to actively be part of the dissemination of content through interaction is one major change in how modern communication works. IS research investigates this phenomenon of users' interaction with content from various perspectives. This chapter sets the determinants for interaction into relation and ultimately shall serve as a basis to derive the research gaps and research questions for this dissertation. Due to the proliferation of misinformation and fake news on social media, a substantial part of the literature investigates interaction based on such concepts (Kim and Dennis 2019; Schuetz et al. 2021).

First, literature provides drivers of interaction based on the individual level. Bessi (2016) takes the perspective of the Big Five personality traits as a basis for social media interaction. The research shows that conspiracy or scientific communities are populated with users having similar personality profiles. The type of interaction is based on simply commenting on posts within the community. This personality profile consists of users who generally enjoy interaction with close friends (low extraversion), are highly emotionally stable, show low agreeableness (i.e., are generally skeptical), have low conscientiousness, and show high openness. Bessi (2016) acknowledges the influence of confirmation bias on interaction but argues that personality traits act as the basis on which individuals act to a great extent. Similarly, Karahanna et al. (2018) argue from a fundamental perspective that interaction on social media is based on basic psychological needs. Among them, they mention the need for autonomy, relatedness, competence, having a place, and self-identity. Beyond those, Lee and Ma (2012) mention the gratifications of information seeking and prior experience with social media for increased interaction. As outlined before, IS research mentions confirmation bias as one of the main influencing factors that determine interaction (Kim and Dennis 2019; Moravec et al. 2019). Lastly, the perception of the source of the content plays a role in whether users interact with the content. Here, perceived credibility was shown to influence interaction (Bapna et al. 2019). This means that if we observe individual differences, research suggests personality traits as a baseline from which psychological needs emerge, which ultimately influence major concepts studied in IS research, such as confirmation bias. The mentioned personality traits and psychological needs already indicate a second category of influence on interaction. Interaction with close friends and relatedness point to social influences as outlined in this chapter beforehand.

The influence of the social aspect is also the second major stream when observing social media interaction in the IS literature. Going further, this category is named group-level influence. Shore et al. (2018) analyze X (formerly Twitter) and find that homophily is a determinant for sharing content. This already indicates how users feel a certain social attachment to a specific group to which they want to feel a social connection. Homophily in the context of fake news sharing showed that it reduces the willingness of users to fact-check and ultimately leads to the further dissemination of fake news (Schuetz et al. 2021). This finding can also be found in similar concepts such as conformity with a community, which has also been shown to influence sharing behavior (Colliander 2019). The mechanisms behind falling in line with a social group are shown to be socializing (Lee and Ma 2012) and also the fear of being

isolated in the case of non-confirmation (Pang et al. 2016). Lastly, research also investigates how such communities or groups are perceived by users. The findings show that only a small portion of users account for public opinion online and therefore act as highly influential users (Ross et al. 2019).

Furthermore, literature investigates the unintended consequences of social media interaction. When talking about unintended consequences, one can distinguish between consequences for the individual and consequences for the group. On the individual level, IS research provides insights into the concept of fear of missing out in the context of social media interaction (Beyens et al. 2016). Fear of missing out, or FOMO, refers to the notion that others may have valuable experiences from which one is absent (Przybylski et al. 2013). This fear ultimately leads to stress within social media users when they are absent from the platform, as they may not be up to date with what is happening with their friends, family, or the world. This means interaction on social media creates a certain dependency relationship with social media. However, on the other side, we observe that too much interaction can also invoke unintended consequences as outlined by Maier et al. (2015). They propose social overload as a concept on social media, as outlined before. On a group level, literature investigates the emergence of echo chambers, meaning communities that curate an environment of like-minded individuals in which opinions that challenge the dominant opinion are rejected or exiled (Shore et al. 2018). And lastly, the spread of misinformation in the guise of fake news is a well-investigated point of contention concerning unintended consequences (e.g., Colliander 2019; Schuetz et al. 2021).

When we look at the literature, we also find different foci on what type of interaction is investigated. As Karahanna et al. (2018) point out, users on social media use the platform to stay in contact with other individuals. However, they do not specify on what metric they are focusing on in terms of liking, sharing, or commenting. A substantial portion of the literature investigates the distinct interaction via sharing of content by users (e.g., Colliander 2019; Shore et al. 2018) or liking (e.g., Beyens et al. 2016; Pang et al. 2016). As pointed out before, sharing and liking introduce a new type of interaction that was introduced in social media platforms, which gives users the opportunity to widely spread content across platforms. Therefore, it is not surprising that research observes such behavior in terms of content interaction, as it provides the basis for phenomena such as fake news or stress.

Study	Context	Individual Level Influence	Group Level Influence	Unintended consequences	Type of Interaction
Bapna et al. (2019)	Investigation of the interaction with the content of brands within social media	Perceived credibility increases the likelihood of interaction.			Like
Bessi (2016)	Influence of personality traits on interaction within Facebook groups	Big-Five personality traits predict interaction in echo chambers		Echo chambers	Comment
Beyens et al. (2016)	Adolescents' use of Facebook with negative ramifications			Fear of missing out is positively related to adolescents' perceived stress related to the use of Facebook.	Like, Reaction
(Colliander 2019)	Fake news sharing on social media		Conformity to others online leads to the sharing of fake news	Sharing of fake news	Share
Karahanna et al. (2018)	Derive affordances from specific social media providers	Social media users are motivated by five psychological needs (autonomy, relatedness, competence, having a place, and self-identity)			Communication
Kim and Dennis (2019)	Evaluation of how presentation format influences the believability of social media posts	Confirmation bias is salient when individuals engage with social media posts.		Belief in fake news	Like, share, comment
Lee and Ma (2012)	Determinants of news sharing behavior on social media	Gratifications of information seeking and prior experience with social media as influencing factors of news sharing	Socializing and status seeking as influencing factors of news sharing		Share
Maier et al. (2015)	Use of social networking sites in conjunction with social overload			Subjective social norm, number of friends, and extent of usage predict social overload.	Extent of usage
Moravec et al. (2019)	Fake News flags and their impact on perceived believability	Confirmation bias strongly influences the believability of social media posts, even in the presence of flags		Belief in fake news	No interaction measured, but belief in fake news is an antecedent of interaction.

Study	Context	Individual Level Influence	Group Level Influence	Unintended consequences	Type of Interaction
Pang et al. (2016)	Interaction with content via comments and click speech on social media		High fear of isolation facilitated self-censorship, while low fear of isolation predicts commenting and liking.		Like, comment
Ross et al. (2019)	Influence of social network actors on interaction		Small numbers of individuals can sway opinions in the online information ecosystem.		Share
Schuetz et al. (2021)	Fake news flags and antecedents for fact-checking		Homophily reduces the act of fact-checking and increases the potential to share fake news.	Sharing of fake news	No interaction measured, but fact-checking behavior
Shore et al. (2018)	Network structure and patterns on Twitter, evaluation of potential echo chambers		Evidence for homophily as an influence on political slant	Echo chambers	Share

Table 1. Literature Review based on the discipline of Information Systems

In summary, the presented papers from the IS literature provide an overview of what we know about influences that facilitate interaction with content on social media from the perspective of a user. Research investigates individual-level and group-level influences, while mainly focusing on the public space of social media. For unintended consequences, the perspective of impact on the whole online information ecosystem is taken, as research investigates the consequences of fake news and the emergence of echo chambers predominantly. This review of the IS literature provides the basis for formulating the research gaps in the following section.

4 Research Questions

In this section the literature review is used to formulate research gaps in order to come up with associated research questions. The research questions serve the purpose to answer the overarching research question and to provide guidance for the further implications for research and practice.

When reviewing the IS literature, there are several studies on influences for interaction (e.g., Kim and Dennis 2019; Pang et al. 2016). However, these instances often focus on a distinct part of the social media information ecosystem. For example, Kim and Dennis (2019) focus on the presentation format of content and how users interact with it. Ross et al. (2019) find that a small portion of social media actors can sway opinions. The question of why such dynamics in communication appear in the first place and how social media has changed the ways in which society communicates nowadays still remains. There is a lack of conceptualization of the communication processes through social media that enable phenomena like fake news, interaction through click speech, or unintended consequences due to social media use. As outlined before, the focus of this dissertation is the perspective of the empowered user. The user, however, now does not stand at the end of a communication process, but further drives it. Therefore, the first research question is:

> ***Research Question 1:*** *How does social media change communication on the level of the user as the audience?*

Past research rarely investigates the various communication configurations that arise in this new complex online information ecosystem. In literature, believability of content is measured based on the content itself, how it is presented, and the source (Kim and Dennis 2019; Moravec et al. 2019). However, the online information ecosystem introduces greater complexity with users being able to share content. This introduces a communication process with content originators and various messengers. Hence, we have insufficient insights into how the process of sharing through various actors alters the perception of content and ultimately, when individuals interact with the content. This is crucial to understand, since social media lives off its users redistributing content, and ultimately, users get to know content through that mechanism. Hence, the following research question emerges:

> ***Research Question 2:*** *How does the mechanism of sharing content through various actors on social media influence interaction by users?*

The literature review shows that the IS literature already has a good understanding of how individual differences influence interaction behavior on social media, especially in terms of confirmation bias (Bessi 2016; Kim and Dennis 2019). Furthermore, we also understand that social influences can drive content interaction (Pang et al. 2016; Shore et al. 2018). However, research rarely takes into account how these two concepts of the individual level and group level interact with each other and ultimately influence interaction. The question remains under which circumstances group-level influence may supersede or diminish in the light of the individual level and vice versa. As there are various individual influences and forms of social influence, this work focuses, among others, on confirmation bias as individual influence, and the mechanism of the spiral of silence theory as a group-level influence. Therefore, the following research question is proposed:

Research Question 3: *How do the individual and the group level influences interact in conjunction to predict interaction on social media?*

As outlined by the literature review, we find several papers that focus on the emergence of echo chambers, meaning spaces of like-minded individuals that reject opinions that antagonize their own (Bessi 2016; Shore et al. 2018). We also find research on group-level influences such as homophily or fear of being isolated (Pang et al. 2016; Schuetz et al. 2021). However, since social media affords the possibility to create distinct spaces, research rarely investigates the concrete circumstances of such spaces. Oftentimes, we find research on the open space of social media, but not much on how such mechanisms like fear of becoming isolated emerge within distinct communities. This is crucial to understand, since, for example, fake news research showed that fake news first emerges in smaller communities, which then burst open and spread the content across platforms (Senaweera et al. 2018). Therefore, the research question follows:

Research Question 4: *What mechanisms lead users to interaction within social media communities?*

Connected with the research question before is also the type of interaction happening within such communities. Research investigates click speech as a distinguishing type of communication on social media (Pang et al. 2016), but rarely focuses on the differences between commenting, liking, or sharing, and under which circumstances different behavior is triggered. Commenting is inherently different from liking since more cognitive effort is needed as users need to consider what to type and actually enact the typing. Liking, as click speech suggests, is only a simple click without much cognitive consideration. Therefore, there is an argument to make that liking is being done much quicker than commenting. Furthermore, literature on the distinct space of social media communities is scarce concerning such differentiation. To understand the differences, the following research question is proposed:

Research Question 5: *How does interaction through click speech and commenting differ in social media communities?*

When interacting in online communities, Maier et al. (2015) find that social overload can be present in such networks. Social overload is an unintended consequence of social media use, as the idea for communities in that regard would be social support that reduces or buffers stress. We have substantial research in the areas of unintended consequences that affect society as a whole, such as the emergence of phenomena such as fake news or the emergence of echo chambers (Bessi 2016; Colliander 2019). However, we are lacking insights into how individuals are affected by the use of social media in terms of the circumstances under which users experience stress. Hence, answering the following research question shall elucidate how stress originates within social media communities on an individual level:

Research Question 6: *What are the reasons for stress while interacting in online social media communities?*

Research provides insight into the dark side of social media in terms of, for example, social overload (Maier et al. 2015) or fear of missing out (Beyens et al. 2016). However, when we look at social media platforms and the negative experiences of users, we still lack an understanding of how these users behave in these cases. This means that such unintended consequences might lead to stress for individual users.

The question then remains about the actual behavioral reaction of users when confronted with such experiences. Since social media provides the opportunity for interaction via comments and click speech, we need to ask how these mechanisms may play a role in how users cope with negative interaction. Hence, the following research question is proposed:

> ***Research Question 7:*** *How do users cope when experiencing negative interaction on social media?*

The seven research questions are settled within the three main chapters of the dissertation. Research question 1 focuses on changes in communication, while research questions 2 and 3 focus on interaction on social media. Research questions 4 and 5 are settled in the subchapter on social media communities. Research questions 6 and 7 contribute to the third chapter of ramifications. The figure below outlines in which chapter the research questions are addressed.

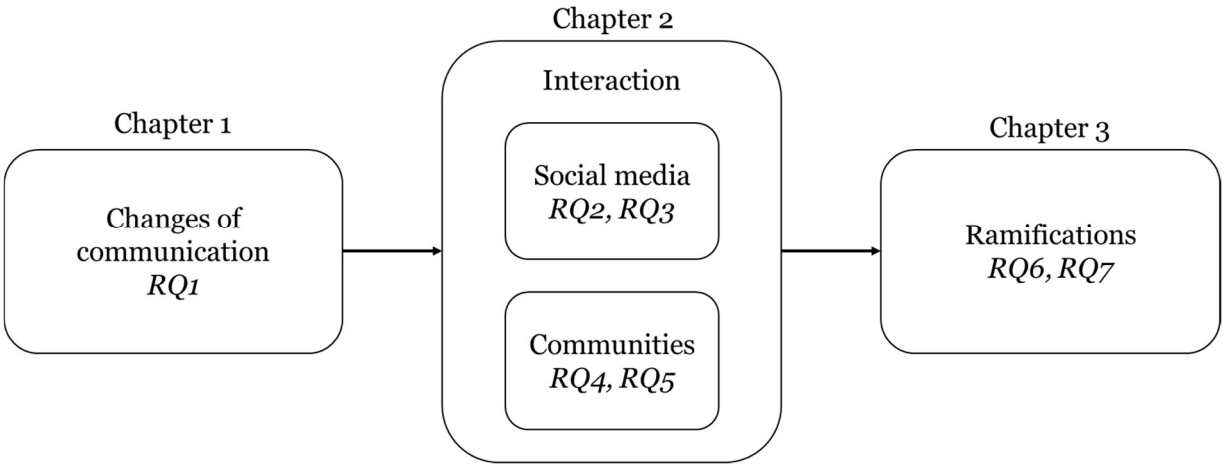


Figure 5. Chapters of the dissertation with the associated research questions

5 Methodology

To answer the proposed research questions, this dissertation applies various research approaches. The identification of research gaps through literature review follows qualitative and quantitative methods to capitalize on the strengths of each individual approach. This section provides an overview of the methodologies that were applied and furthermore elaborates on the specifics of each approach. The section is structured as follows: first, there is a deep dive into how literature reviews were conducted with their particularities. Then, the section provides an overview of the applied qualitative research. Lastly, quantitative research is elaborated on, and how this approach was found applicable in the dissertation.

5.1 Literature Review

This dissertation contains two literature reviews. First is the literature review in the introductory paper. The goal of this review was to identify research gaps within the communication on social media platforms. Second, the literature review in **Paper I** addresses specifically how communication changed with the introduction of social media platforms. The second literature review specifically dives into changing roles of communication through social media to explain the emergence of the phenomenon of fake news. Traditionally, news media organizations had the means to reach big target audiences (Westley and MacLean Jr 1957; White 1950). This gatekeeping characteristic was essentially transferred to social media platforms as intermediaries of information (Bro and Wallberg 2014). To understand the new communication patterns, one has to understand the changing roles within such an online information ecosystem. Based on the findings of the second literature review, the dissertation can better contextualize the individual contributions based on the different communication roles and patterns, and provides ramifications of such changes as seen in the emergence of the phenomenon of fake news.

The two literature reviews used different approaches to make sense of the literature. Literature review one contains the proposed steps by Webster and Watson (2002). In the first step, the scope of the literature review was set, which included the research context and IS as the field of research. A keyword search was performed using titles, abstracts, and keywords such as “social media” and “social network*”. The scope of the search contained leading A and B-rated IS journals. In the second step, the articles were scanned for the focus of the literature review: determinants and unintended consequences of social media interaction with content. Here, the codes of individual level and group level influence emerged, as well as the type of interaction observed. Finally, the papers were analyzed and set in context to show how each of them contributes to a better understanding of social media interaction among users.

The second literature review (**Paper I**) proposes a grounded theory approach to literature reviews. Here, the grounded theory literature approach, according to Wolfswinkel et al. (2013), was conducted. In general, grounded theory is an approach with the goal of generating theory based on data that is systematically collected (Strauss and Corbin 1990). The building of theory through literature involves five steps according to Wolfswinkel et al. (2013): *define, search, select, analyze,* and *present*. In their paper, Wolfswinkel et al. (2013) describe that the methodology is flexible in the sense that it can be

adapted to the specific needs and goals of the literature review. In this case, communication theories are already outlined in the literature. This is why the starting point for the change in communication roles and patterns was past roles and patterns, according to Westley and MacLean Jr (1957). With the basis of traditional communication processes, it is possible to *define* the research area of interest and to provide inclusion and exclusion criteria. The aim was to investigate the social media platforms and the phenomenon of fake news. Accordingly, a broad literature *search* on the Web of Science was conducted with the keywords "social media" OR "social network*" AND "fake news", as fake news served as the case for changing communication roles and patterns. In the *select* phase, articles were only further investigated if they provided meaningful insights into the communication process. In the analysis step, the remaining papers were coded, and the communication model was refined in several iterations. Lastly, the literature results could be *presented* through the refined communication model, and future venues for research could be identified.

5.2 Qualitative Research

In cases in which research was still relatively immature, qualitative research methods provide an ideal approach in order to investigate a phenomenon initially. Qualitative research shows high efficacy if there is an insufficient understanding of the relationship between context and phenomenon or behavior (Yin 2009). This approach was used in **Paper IX** in order to investigate distinct strategies of content creators on social media, specifically on Instagram. Since literature has not yet provided clear coping strategies for content creators, the qualitative approach in the form of semi-structured interviews (Adams, 2015) showed great potential in extracting behavior patterns of content creators after experiencing a negative interaction on the platform. **Paper IX** includes semi-structured interviews from 19 Instagram content creators. One has to clarify that these content creators do not use Instagram as a source of income, but show a consistent history of posting behavior. The semi-structured interviews first focused on what exactly constitutes a negative interaction on the platform, and second, how content creators react after they experience such a situation. Hence, the semi-structured interview guideline incorporated open-ended questions to make sure different viewpoints could be examined (Yin 2009). Each interview was transcribed and coded with MAXQDA 12. One researcher coded the transcripts, while another cross-checked the findings. Disagreements within the coding process were resolved in a cooperative manner through in-depth discussions. Most disputes arose from the labelling of the individual coping strategies, as their label should incorporate the nature of the behavior that is most ideal. Since we started from the coping literature (Folkman & Lazarus, 1985), the first codes were deductive in nature. These were stressors, resources, and coping concerning emotion and problem-focused approaches. After this initial coding process, we applied inductive coding to come up with sub-codes of the main codes. Therefore, we gained distinct coping strategies under the main code of coping and categorized them between emotion-focused and problem-focused coping strategies. The findings on the coping strategies and origins of stress for content creators were then used in a follow-up qualitative study in **Paper X**.

5.3 Quantitative Research

In this dissertation, **Papers II, III, IV, V, VI, VII, VIII,** and **X** build on quantitative research through surveys and experimental designs. In the following section, different approaches to quantitative research are outlined and categorized by survey and experimental setups.

5.3.1 Experimental and Survey Research

The main body of research in this dissertation builds on self-reported data from surveys and experimental research. Even though survey research can be biased, due to phenomena such as social desirability bias (Krumpal 2013), survey research provides the opportunity to specifically investigate the relationship of the constructs in question. Furthermore, it is possible to control for special groups to make sure that the findings can be contextualized.

As this dissertation focuses on the behavior of users on social media in conjunction with interaction and information, there are two major dependent variables: believability and opinion disclosure.

Study 1

A 2x2x2x3 factorial between-subjects study design was developed to investigate the opinion-polarizing effects of three characteristics of online network communication, based on the perceived similarity with the actor sharing fake news on Facebook. The characteristics of the actor and the communication process in which fake news is shared are manipulated by creating mockup Facebook profiles and pages. The experiment builds on the similarity-attraction paradigm (Ruijten 2021) through different actors who share headlines. This experiment manipulated the similarity between the study participants and the Facebook profiles that were developed. The similarity, thereby, is based on the political affiliation.

Facebook allows for the creation of two types of sites: organizational pages to connect with target groups or individual user profiles, for every ordinary private user. Such sites are created for each actor in the study. Regarding the actor characteristics, political (similar, dissimilar) bias is controlled by manipulating the content on the Facebook pages for profiles with news-media-like presentation and advocates (i.e., influential personalities), as well as Facebook profiles for users. These provide clues of Democratic or Republican party affiliation, exposing participants to either politically similar or dissimilar actors. The account type (news media, politician, user) content varies based on the avatar and the type of Facebook site: news media pages are patterned after news websites, politician pages are characterized as public political figures (representatives), and user profiles indicate private users.

During the communicative process, mediation (direct, mediated) is manipulated by exposing participants to fake news posted directly by an actor or shared by a second actor. In the mediated condition, actor roles are distinguished by whether the post was originally created by the actor (source) or shared by the actor (messenger).

To establish external validity, each Facebook site is conceptually based on a randomly drawn real-world example and altered to protect personal privacy while allowing for experimental manipulations. All

names, avatars, jobs, friends, and images were changed to ensure anonymity. Structural page information, including likes, shares, and non-related posts, was maintained to evoke perceived authenticity. Overall, 1.232 participants were recruited to rate headlines according to their believability. **Paper III** and **Paper IV** build on the data set of the study with a different focus. **Paper III** uses the whole data set and focuses on how similarity and dissimilarity through different actors influence the believability of the headlines.

Paper IV uses a subset of 931 surveys to evaluate the influence of the messenger since direct effects were also measured in the main study. It focuses on believability as the dependent variable. The paper evaluates how users perceive false information through messengers and how these finally influence the believability of news-like media. Since confirmation bias has been shown to be a driving force in how social media users believe information (Kim and Dennis 2019), the study gathered data around how the relationship between confirmation bias and believability is altered by messenger effects.

Study 2

A social media scenario was constructed to test the hypotheses. In this scenario, the participant is suggested to be on a social media website where commenting and liking posts are possible. Participants provided information about their most frequently used social media platform, which they were asked to recall when engaging with the scenario. The participants engaged with a post about COVID-19 vaccination, reaffirming previously held beliefs. This scenario was designed to evoke confirmation bias.

After reading the scenario, participants rated how likely they were to comment on or like the post. No specific opinion was imposed on participants, allowing them to imagine an opinion they already agreed with, thereby placing themselves in a state of confirmation bias.

In the main study, participants reported their trust and risk beliefs concerning privacy on social media before being introduced to the scenario. At the end of the main study, participants were asked to report how important they found the topic and to identify the topic from a list of options, including the COVID-19 vaccine. Responses that did not pass this attention check were excluded from the evaluation. In the post-study, demographics such as age, gender, and education were collected. A total of 168 participants were recruited, and 148 passed the attention check. **Paper V** builds on this dataset and, thereby, investigates commenting and liking on social media alike.

Study 3

Study 3 focused on the spiral of silence theory as the theoretical basis for opinion disclosure. The experiment used three scenarios that focused on prominent topics in the mainstream US media landscape during 2020, after which the experiment was conducted: immigration, the presidential election, and COVID-19. These topics were selected because the sample consisted of American citizens, and the chosen topics were of equal importance in public opinion during that year (Hrynowski 2020). Scenarios from previous spiral of silence research were adapted to a social media context (Neuwirth et al. 2007; Scheufele et al. 2001). To ensure the relevance of the issues, participants rated their importance on a Likert scale ranging from "not at all important" (1) to "extremely important" (7). The average

importance scores were 5.012 for immigration (standard deviation (SD): 1.410), 5.232 for the presidential election (SD: 1.509), and 5.287 for COVID-19 (SD: 1.422).

Paper VI focused on trait-based and state-based fear of isolation as the mechanism of the spiral of silence theory. The scenarios were tailored to participants in a specific social media community, emphasizing the dominant community opinion on each issue rather than focusing on the messenger of the opinion, who was not identified, to eliminate the influence of social ties. In each scenario, participants indicated how likely they would be to share an opinion in the situation, which served as the dependent variable. In **Paper VII**, the focus lies on the state-based fear of isolation in conjunction with community identity and how both influence opinion disclosure, specifically targeted at clickbait speech.

This setup was chosen because it provides a controlled environment to specifically test the influence of fear of isolation. Although social media platforms offer content that reflects authentic opinion disclosure, it would be difficult to determine whether such disclosure was genuinely motivated by fear of isolation or other factors. Therefore, a scenario-based study allows for better control over these influences and provides a clearer basis for attributing opinion disclosure to fear of isolation.

Study 4

In study 4, participants were instructed to imagine posting an Instagram photo, which was introduced at the beginning and served as an anchor (Kahneman et al. 1982) for likes. This initial anchor provided a foundation for subsequent manipulation. Due to the initial qualitative research in **Paper IX**, a non-commercial Instagram user typically received an average of 234 likes. Participants were then asked to share their perception of obtaining this number of likes. The results indicated that all participants were content with the number of likes received in the anchor treatment.

After viewing the anchor post, participants were told to envision posting a second one, which received only a third of the initial likes and included a negative comment, following the approach used by Stsiampkouskaya et al. (2021). This stage aimed to manipulate their emotional response. Participants were then prompted to indicate their feelings (anger or sadness) and describe their coping strategies during individual treatments. Anger and sadness were used to represent high (anger) and low (sadness) arousal emotions, allowing for an exploration of how these different emotional states influence coping strategies. Finally, participants evaluated their positive outcomes after implementing their coping strategies.

In the treatment, to ensure greater authenticity, male participants were shown photos of a male individual, while female participants received photos of a female individual. For those who identified as neither male nor female, they were randomly assigned to one of the two treatments. This approach aimed to enhance the likelihood that participants could relate to the scenario, which was crucial as they needed to imagine themselves posting Instagram photos.

An “explorer-type” image was chosen for the treatments. This type of image, which features a person paired with a scenic background, is among the most commonly posted on Instagram and is associated with popular hashtags such as #nature, #travel, #photography, and #me (Top-Hashtags 2023).

Paper X is based on the dataset of study 4 and investigates distinct coping strategies. These broadly fall under the category of opinion disclosure but are not exclusively part of the category. As outlined beforehand, the coping strategies could be investigated through a qualitative study. Through the quantitative approach it was possible to validate which coping strategies occur and lastly which efficacy each of those showed in the end. The study contained 139 completed surveys.

Study 5

This dissertation contains two studies that were purely based on a survey without any treatment associated with it. This survey aimed for a general perception of news on social media, specifically X (formerly Twitter) and the likelihood of news sharing across the platform. Hence, the news sharing attitude was the depended variable in the survey. Paper (Endorse) uses the data set of the study and investigates the influence of information verification and source credibility on the news sharing attitude.

Study 6

The second survey focused on the relationship between social support and social overload in the context of online mental health communities on social media. The survey contained participants actively participating in said communities. In total 110 full responses could be gathered. Paper VIII uses this dataset to determine how cohesiveness and universality influences social support and in return how social support relates then to social overload.

5.3.2 Validity

All data sets aimed for a balanced sample in terms of gender. Furthermore, multiple attention checks were implemented to ensure that participants read the survey. The attention checks were conducted in two ways. First, questions were asked to be answered in a specific manner (e.g., click “strongly disagree”). Second, after receiving a treatment, the survey asked participants about the content of the treatment (e.g., after showing a scenario, participants were asked about what the scenario was about). All surveys that did not pass the attention check were omitted and were not part of the evaluation of the results. The following table shows an overview of the different surveys.

Study	Research Context	Research Objective	Dependent Variable	Participants	Papers
1	Facebook	Investigating determinants of believability of news-like content	Believability	931	III, IV
2	Social Media	Investigation of privacy-related concepts to opinion disclosure on social media	Opinion disclosure (commenting, liking)	148	V
3	Social Media Communities	Application of the spiral of silence theory within social media communities	Opinion disclosure (commenting, linking, sharing)	164	VI, VII
4	Instagram	Testing the relationship between negative emotions and coping strategies after a negative interaction	Coping strategies (content deletion, acting out of spite, defense, blocking)	139	X
5	X (formerly Twitter)	Evaluating how the general perception of the platform influences news sharing	Opinion disclosure (sharing)	148	II
6	Online Mental Health Communities on Social Media	Contextualizing social support and social overload within social media communities	Social overload	110	VIII

Table 2. Quantitative data collection through quantitative research

The data were collected using Amazon Mechanical Turk (mTurk) for all studies. In that regard, research provides guidance on how to set up mTurk to acquire reliable data. Hence, the papers follow the suggestion of Lowry et al. (2016), who propose mTurk participants with a master's qualification. These workers have a generally high approval rating of tasks and have shown a multitude of reliable data in the past. Furthermore, with mTurk, it was possible to directly target groups, as it was needed for several studies. Where a master's qualification was not applicable, the minimum approval rating of workers was 95%, and a minimum of 500 tasks completed. For example, for study 4, only Instagram content creators were targeted, as the data from a user who only consumes content on social media would not provide reliable data on coping strategies after having a negative interaction associated with their own content.

5.3.3 Card Sorting

Every study builds on pre-tested and valid instruments to measure the respective constructs within the research models, except for the construct of the state-based fear of isolation in study 3. The distinction between trait-based and state-based fear of isolation is not as clear in the literature, and therefore, there are only rudimentary instruments that investigate the phenomenon. Hence, prior to study 3, new items were developed for the state-based fear of isolation based on the characteristics that we know from the literature, and finally through card-sorting.

To develop the new scale accurately (Moore and Benbasat 1991), a card-sorting procedure with four individuals was conducted to test the validity of the five items. Beforehand, similar concepts to fear of isolation were identified from literature to provide a solid case that the newly developed instrument reflects the state-based fear of isolation even in proximity to similar constructs. The related concepts were anxiety and fear of negative evaluation, which refer to the emotional state of fear or anxiety in a particular context. Five items for the state-based fear of isolation (self-developed), three items for anxiety (Venkatesh et al. 2003), and five items for fear of negative evaluation (Watson and Friend 1969) were chosen for the card sorting test. The items were tailored to the social media context to prevent participants from recognizing which items belonged together. The card-sorting ultimately showed that the state-based fear of isolation items were correctly mapped (correct classification: 85%). Single items were only misclassified once. Based on metrics of interrater reliability (Cicchetti and Sparrow 1981; Landis and Koch 1977), state-based fear of isolation is perceived differently and has distinct characteristics that distinguish it from related concepts.

5.3.4 Structural Equation Modeling

The majority of the data sets within this dissertation were evaluated by performing structural equation modeling (SEM). SEM belongs to the class of multivariate statistical tools to assess research models and, therefore, hypotheses in various scientific disciplines such as social sciences, health, or managerial research (Bagozzi and Yi 2012). Structural equation modeling consists of two parts. First is the measurement model, which makes it possible to assess the relationship between the measurements and the latent variables. Second, the assessment of the relationships between these variables is called the structural model (Hair Jr et al. 2014). As a statistical tool, this dissertation mainly uses SmartPLS and, therefore, SEM with the partial least squares method.

To assess the validity of the structural equation modeling, various quality criteria are applied first in the measurement model (Hair Jr et al. 2014) for reflective measurements. First, internal consistency reliability is evaluated. Traditionally, Cronbach's alpha estimates the reliability based on the intercorrelations of the observed indicator variables. However, Cronbach's alpha is sensitive to the number of indicators and generally underestimates the internal consistency reliability. Therefore, a different measure is considered: composite reliability. The measures for internal consistency reliability vary between 0 and 1. Composite reliability measures between 0.70 and 0.90 are desirable (the same applies to Cronbach's alpha). Values above 0.95 are not desirable as they indicate that the items measure the same phenomenon, and therefore are too similar in nature (Hair Jr et al. 2014). Lastly, composite

reliability can be used as the upper boundary of internal consistency reliability, as in contrast to Cronbach's alpha, it overestimates the internal consistency reliability (Hair Jr et al. 2014). Hence, both quality criteria are considered.

Second, convergent validity is assessed, which refers to the extent to which a measure correlates positively with alternative measures of the same construct (Hair Jr et al. 2014). To assess convergent validity, researchers consider outer loadings and average variance extracted (AVE). For the outer loadings, high loadings are desirable, generally above the threshold of 0.708, with a significant relationship with the construct. Furthermore, the AVE refers to the grand mean value of the squared loadings. This means all loadings squared and summed up are divided by the number of indicators. Values above 0.50 indicate convergent validity, as on average the construct explains more than 50% of the variance of its indicators (Hair Jr et al. 2014).

Internal consistency reliability
Cronbach's alpha and composite reliability should be above 0.70, whereas Cronbach's alpha serves as the lower bound and composite reliability as the upper bound.
Convergent validity
Indicator outer loadings should be above 0.708 and significant for their respective construct. The AVE should be higher than 0.50
Discriminant validity
Indicator loadings should be highest for their respective construct (cross-loadings)
Fornell-Larcker criterion: AVE larger than the squared correlation with any other construct
HTMT value below 0.90

Table 3. Quality criteria thresholds overview for structural equation modelling using the partial least squares method

Third, discriminant validity is assessed. The assessment refers to the extent to which a construct is distinct from other constructs in the model (Hair Jr et al. 2014). For example, there can be a highly significant relationship between two constructs, but only because the two constructs measure essentially the same. In this situation, the insights into the relationship are null. For the discriminant validity, there are generally three ways to be assessed: cross-loadings, the Fornell-Larcker criterion, and the Heterotrait-monotrait (HTMT) ratio (Hair Jr et al. 2014). Cross-loadings refer to the indicators. In this case, the indicators should load highest with their respective constructs in order to conclude discriminant validity. The Fornell-Larcker criterion compares the square root of the AVE values with the variable correlations. The AVE needs to be larger than the squared correlations with other constructs in order to assess discriminant validity (Hair Jr et al. 2014). Lastly, the HTMT ratio is the ratio between-trait correlations to the within-trait correlations. Values of the HTMT ratio above 0.9 indicate a lack of discriminant validity (Henseler et al. 2015). In the following table, an overview of the quality criteria for the reflective measurement model is presented.

For formative measurement models, there are different criteria to apply, namely the weights and multicollinearity. The indicator weights should be significant, and multicollinearity, meaning that two variables are highly correlated, should not be given. Generally, the variance inflation factor (VIF) is used

to detect multicollinearity. VIF values should be below 5 in this regard (Rogerson 2019). The following table shows the quality criteria for the formative measurement model.

Weights
Indicator weights, absolute impact, and significance
Multicollinearity
VIF values below 5

Table 4. Quality criteria for the formative measurement model

Magnitude of path coefficients (β)
Between 0 and 1, the closer the value to 0, the weaker the relationship
Coefficient of determination (R^2)
Rule of thumb interpretation > 19% = weak > 33% = moderate > 67% = substantial
Effect size (f^2)
> 0.02 = small effect > 0.15 = medium effect > 0.35 = large effect

Table 5. Path coefficient, R^2, and effect size values and their interpretation

After the assessment of the measurement model, the relationships between the constructs can be assessed. There are various values to consider when assessing the relationships, mainly: path coefficients, significances, R^2, and effect sizes. Path coefficients refer to linear regression weights and are used to assess the statistical relationship between constructs. Path coefficients range from 0 to 1, while higher values suggest a stronger effect. Significances (p) show whether the influence between constructs actually holds and the probability of rejecting the null hypothesis. Meaning, based on a random sample from a data set, with $p < 0.05$, there is a 95% probability that the null hypothesis will be rejected, if the relationship is significant at the 0.05 level. R^2 refers to the coefficient of determination. It shows the percentage of variance that can be explained by the proposed model. Lastly, the effect size (f^2) is a measure of the actual effect of the variables on a construct (Hair et al. 2012; Hair Jr et al. 2014). The following table provides an overview of the interpretation of the distinct values.

5.3.5 Moderation and Mediation

The analysis of relationships in structural equation modeling is not only bound to one variable influencing another. Influences can shift strongly based on a third variable. If a third variable influences relationships, this construct is called a moderator (Baron and Kenny 1986; Hair et al. 2012; Hair Jr et al. 2014).

Moderation variables can change the strength or even the direction of the relationship. In the analysis, an interaction term between the predictor and the moderator is generated and tested to determine whether the interaction term significantly influences the dependent variable (Baron and Kenny 1986).

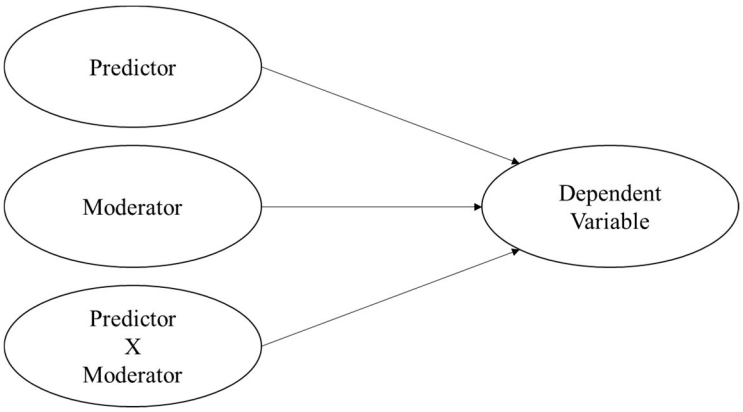


Figure 6. Modelling of moderation effect within structural equation modelling based on Baron and Kenny (1986)

In structural equation modeling, the moderating effect is usually depicted with the moderator influencing the relationship between two constructs. Hence, the moderator's arrow targets the relationship, as shown below.

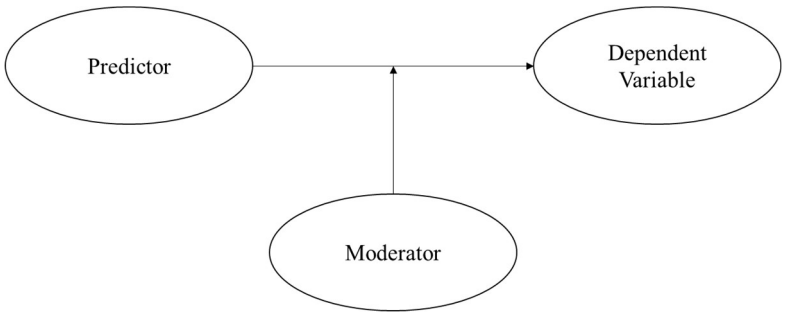


Figure 7. Representation of the moderation effect in research models

An example of moderation could be the relationship between sleeping hours and attention span, moderated by coffee consumption. We would hypothesize that the longer the sleep, the higher the attention span throughout the day. However, if we have individuals consuming coffee on top, we would expect that the interaction term has an even stronger effect on the attention span. In this case, coffee consumption would positively moderate the influence of sleeping hours on attention span throughout the day.

Besides moderation, there is also mediation. Meditation refers to independent variables influencing a dependent variable through a third variable (Baron and Kenny 1986). Similar to the moderation, we assume a research model with three variables. There are three conditions under which a variables assumes to be a mediator (Baron and Kenny 1986): (1) variations in the independent variable significantly accounts for variations in the mediator, (2) variations in the mediator significantly influence variations in the dependent variable, (3) in the strongest case the direct effect of the

independent variable on the dependent variable becomes insignificant with the presence of the mediator.

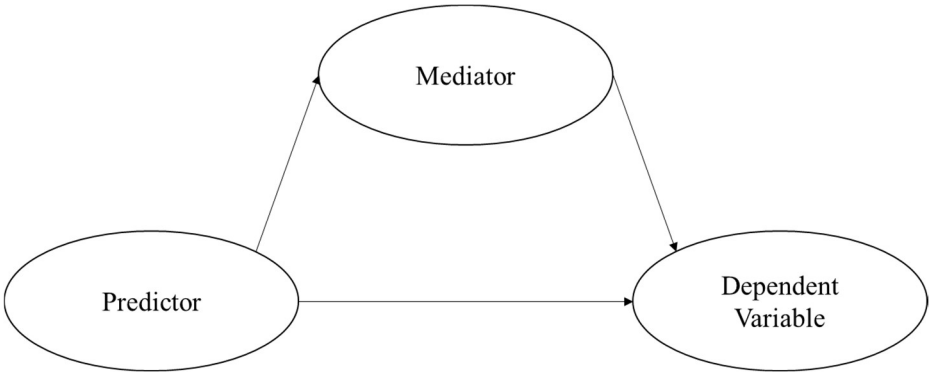


Figure 8. Representation of mediation based on Baron and Kenny (1986)

For example, we assume we want to investigate the relationship between hot temperatures and drowning incidents. In our statistical analysis, we may find that higher temperature positively influences drowning incidents. However, when introducing the number of people going to the beach as a mediator, we may find that hot temperatures significantly influence the number of people going to the beach. Hence, if more people go to the beach, we statistically expect more drowning incidents. Hence, the number of people going to the beach acts as a mediator for hot temperatures and drowning incidents.

5.3.6 Summary

The dissertation applies various methods in order to investigate the overall research questions. The detailed approach is always laid out in the different papers. This section provides a basis to understand the main parts of the methodology and the analysis. The following section presents the main results aggregated based on the individual papers in this dissertation.

6 Main Research Results

This dissertation contains ten papers as outlined in the introduction. Each paper contributes to the proposed research questions in different ways. The following table provides an overview of the results of the papers to the different research questions.

	Empowerment of users as active audience. Information customization through filters.	Social media's inherent mechanisms, such as commenting and sharing, elevate users to active content creators and distributors.	Interaction is based on individual-level (mainly confirmation bias) and group-level (social ties) influences that interact with each other.				
	The responsibility of information curation and verification is now heavily skewed towards the audience.						
	Users receive content from various different account types, such as opinion leaders, news media, and other users.	Communication on social media is either mediated or immediate, based on the sharing mechanism.					
	Users are able to trace back information to its origin directly.	Conceptualization of account types; content is funneled through accounts until it reaches users and influences the perception of content.	The interaction between confirmation bias and how users perceive differently based on account types.				
	Interaction on social media becomes salient and visible to other participants on the platform which opens up privacy vulnerabilities for users.		The personal importance of a topic heavily favors users' behavior in commenting on content, in contrast to liking it.		Topic importance directly influences commenting behavior but does not significantly influence click speech.		
	Highlighting the importance of social media community members in how communities are perceived through the comments of the members.		State-based fear of isolation leads to opinion disclosure against the own opinion, or self-censorship in order to remain within the social media community. The state-based concept mediates the trait-based fear of isolation, which does not have a direct effect.	Identification of different types of fear of isolation as a driving force for self-censorship or opinion disclosure against one's own beliefs.	Fear of isolation leads to commenting against the own opinion in specific contexts. If topic importance is high, the relationship is strengthened.		

	Emergence of communities on social media platforms with distinct topics and associated communities with individual rules and beliefs.		Community identity as an influence on state-based fear of isolation	Community identity and fear of isolation as drivers of user interaction when confronted with content that opposes personal beliefs. Group-level influences overcome confirmation bias.	Click speech is an easy-to-enact mechanism that leads to opinion disclosure against one's own convictions.		
				Cohesiveness and universality influence social support, which in return leads to user interaction in an act to give something back. Giving too much back in return fosters social overload, which inhibits interaction.		Identification of social overload through social support as the origin of stress within social media communities.	
						Identification of insufficient likes and negative comments as negative user interaction on social media.	Identification of problem-focused coping strategies on social media and offline.
						Testing and validating insufficient likes and negative comments as origins of negative emotions, which cause stress.	Coping strategies that facilitate further content creation can reduce sadness. Destructive coping strategies cannot reduce anger on social media.

Table 6. Overview of papers and their contribution to the individual research questions

Paper I: Information Flow, Gatekeeping, and Fake News: Towards an Integrative Model for Social Media Communication

The first paper of this dissertation contains a literature review to understand the changes in communication through social media platforms. Social media fundamentally changed communication roles and patterns, and specifically empowered individual users through sharing mechanisms. The paper starts from a traditional communication perspective in which news media organizations are assumed to be the gatekeepers of information, while the audience is a passive consumer, as indicated in the communication model of Westley and MacLean Jr (1957).

By investigating the Fake News literature, several changes could be observed in such a traditional understanding of the communication process. First, we see a change within the role of the gatekeeper. While traditionally, news media organizations assumed this role, now, social media platforms hold the role of the information gatekeeper. Also, platforms are labelled as intermediaries since their main purpose is to connect the actors on the platform. Second, news media organizations are losing power in this process, as they assume the role of opinion leaders, which builds on their reputation as a determining factor for their success. Third, with users being able to share content, they assume an active part in the communication process, which does not render them solely passive consumers of content from opinion leaders. The following figure illustrates the communication roles and patterns, as well as the findings for the context of the Fake News literature.

In summary, the paper elaborates on the changes in communication roles and patterns. Among them is the changing role of the audience when interacting in this online information ecosystem. Hence, the paper provides answers for RQ 1, 2, and 3. **Paper I** explicitly elaborates on the empowerment of users as the audience through social media's inherent mechanisms, such as liking, commenting, and sharing. Also, the possibility for users to customize their information diet through filter mechanisms is elaborated on (RQ1). Furthermore, the paper presents the ramifications of such interaction through liking, commenting, and sharing. The information flow through an intermediary is a key aspect of the literature review, and it shapes how audiences engage with content on social media (RQ2). Lastly, the paper identifies factors leading to interaction. Among them are individual and group-based influences that, in turn, facilitate phenomena such as Fake News dissemination (RQ3).

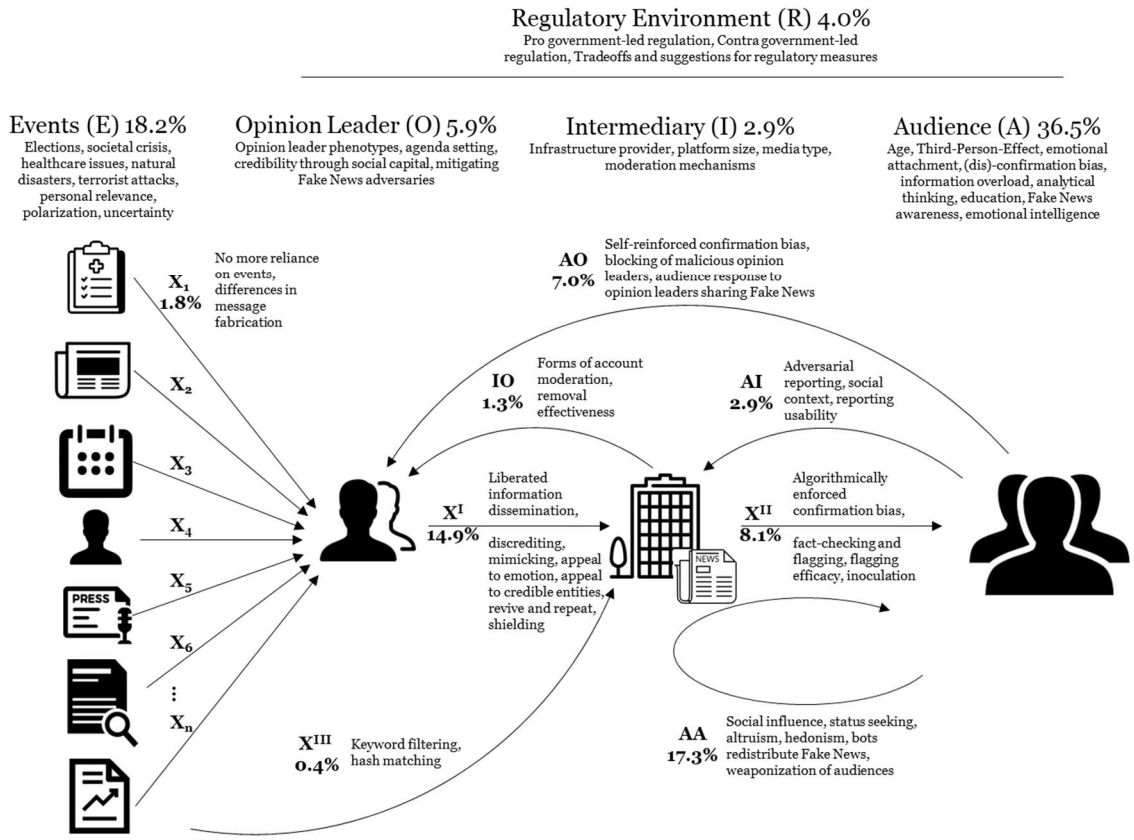


Figure 9. Results from the literature review on Fake News with communication roles and patterns

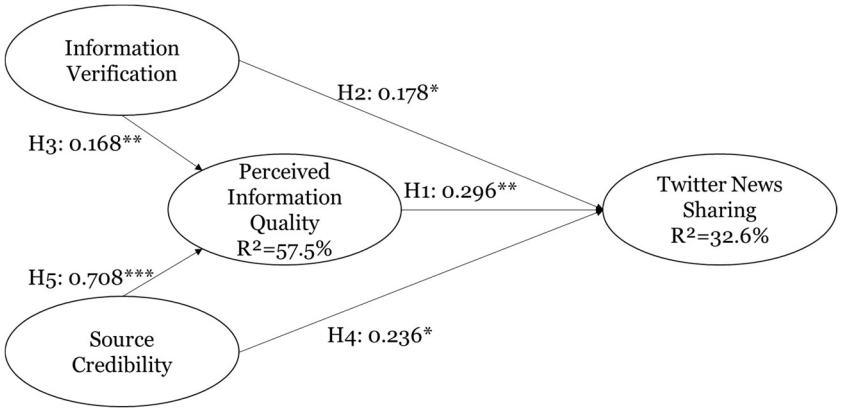
Paper II: Endorse the Source – The Impact of Information Assessment on News Sharing Behavior[1]

The second paper takes a general approach to sharing news on social media, specifically X (formerly Twitter). The objective of the paper is to elaborate on how users on Twitter perceive information quality, with a focus on whether they turn to the source as a heuristic to determine quality, or if they rely on their own verification skills.

The results show that source credibility is the dominant factor for the assessment of perceived information quality. Perceived information quality, in return, significantly influences the general sharing attitude on the social media platform. This means that users generally rather rely on heuristics to determine information quality and do not want to take the extra step to verify information on their

[1] Haug, Maximilian (2019) Endorse the Source – The Impact of Information Assessment on News Sharing Behavior. In: (Proceedings of the) 25th Americas Conference on Information Systems (AMCIS), August 15-17, 2019, Cancun, Mexico.

own. Since social media floods user feeds with large amounts of information, it is not surprising that users turn to easy-to-process heuristics, such as determining who is trustworthy to share content.



Note: p-values. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.09$, ns = not significant.

Figure 10. Determinants of Twitter News Sharing with information verification, perceived information quality, and source credibility

In summary, **Paper II** mainly contributes to answering RQ1. Due to the fact that news media organizations lost power in their gatekeeping role, social media alters communication in the sense that the responsibility for the evaluation of content is shifted more towards users. They do so by redefining their own information ecosystem. Users curate their own information ecosystem through what content they consume (algorithm-driven) and what accounts they personally follow (user-driven), rather than passively receiving information curated by traditional media gatekeepers. Lastly, sharing driven by users is presented as a key behavior shaping how information circulates. This demonstrates that communication on social media is not just about consuming information but also about engaging with it, endorsing it, and influencing others on a user basis.

Paper III: The New Media(tion): The Impact of Social Media Communication Characteristics on the Believability of News

Paper III takes different perspectives on how users, as the audience, engage with news and their believability. In this paper, the relationship between perceived similarity and news believability was tested. Perceived similarity was also established through manipulations of political affiliation. Different account types were observed, such as news media, politicians, and users as the audience, and their different perception that leads to different believability of news. Also, the difference between immediate communication, meaning users directly interact with original content, and mediated communication, meaning users interact with a messenger sharing original content, is tested. In total, four hypotheses were tested.

First, the main relationship between perceived similarity and news believability is established. However, this relationship is subject to intricate moderation effects as outlined below. Second, opinion leaders such as politicians increase the effect of similarity and believability. Third, there was no exacerbation of

the relationship through the messenger, but messengers and sources have independent main effects. Beyond these findings, a mixed-amplification effect was found in which communication through a dissimilar and similar actor in conjunction increases believability most. Lastly, mediated communication through different actors exacerbates the relationships of perceived similarity on believability compared to immediate communication. The following three figures outline the differences between the moderation effects.

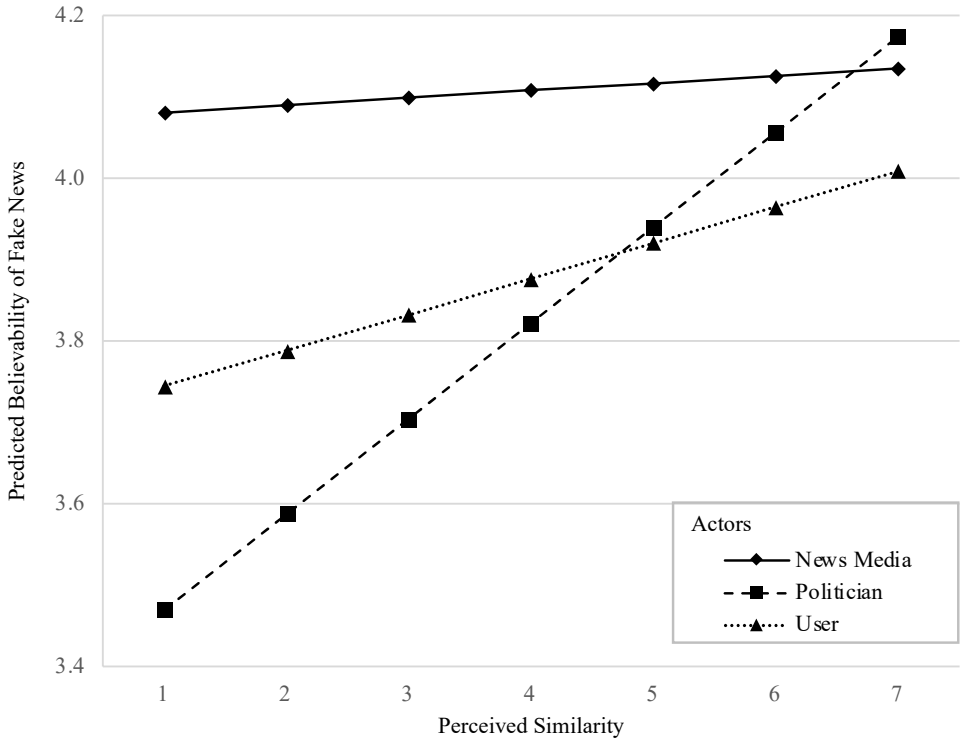


Figure 11. Interaction effects of actors' predetermined roles (i.e., news media, politician, user) with perceived similarity on news believability.

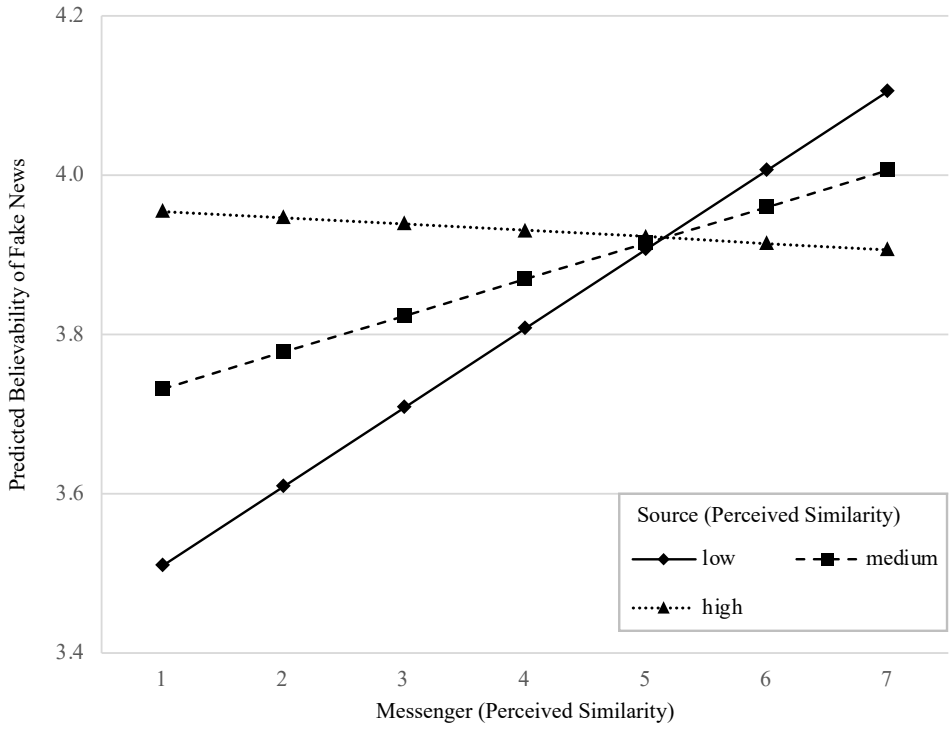


Figure 12. Interaction effects of the perceived similarity with actors of different emerging roles (i.e., messenger, source) on news believability.

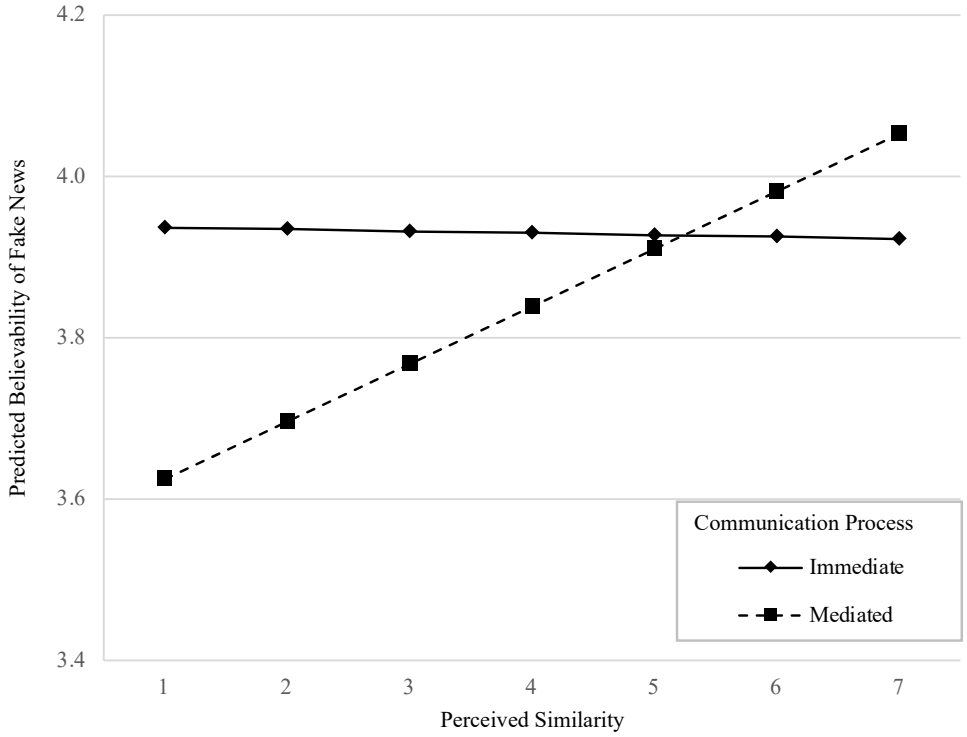


Figure 13. Interaction effects of the communication process (i.e., mediated, immediate) with perceived similarity on news believability.

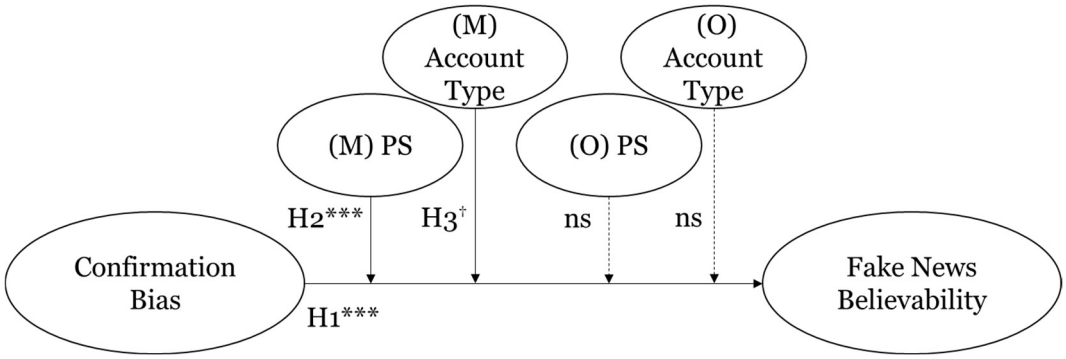
In summary, the paper taps into different occasions on which users engage with content on social media. Users interact with content from various account types, and among them are also other users who are not necessarily distinct opinion leaders or news media organizations. Hence, on a broader scale, with the assessment of the believability of content, social media's distinct behavior, such as liking,

commenting, and sharing, is associated with it (RQ1). Furthermore, the paper specifically investigates the differences between immediate and mediated communication. This form of communication is facilitated by the sharing mechanism and therefore provides further answers for RQ2.

Paper IV: Beware of the Messenger! How Social Media Account Characteristics Moderate the Confirmation Bias

The fourth paper investigates how the mechanism of sharing ultimately influences how users, as the audience, view information. Since sharing provides the opportunity to further disseminate social media posts from others, the question remains how the different actors influence such information flow. For example, a post can be shared by another actor before a user reads the post. The question remains which account, the originator of the post or the messenger of the post, influences the believability of the post in this chain of sharing.

Therefore, **Paper IV** utilizes the dataset of study 1, in which false headlines of Facebook posts were funneled through different actors. The paper builds on perceived similarity with the originator and messenger accounts and how it moderates confirmation bias. The results show that especially the similarity (established through political affiliation) towards the messenger has a significant influence on the relationship between confirmation bias and the believability of the news headline. In particular, the similarity exacerbates the relationship between the two concepts. Furthermore, account types were investigated, such as politicians, news media organizations, and users within the audience who share posts. The results show that opinion leaders without an attachment to news media also strengthen the influence, in contrast to news media accounts.



Note: PS = perceived similarity; (M) = messenger; (O) = originator; p-values. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.09$, ns = not significant (two-tailed significance).

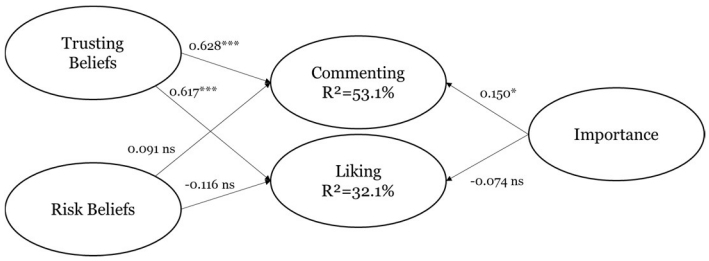
Figure 14. Research Model Results: Perceived Similarity and Account Type of Messengers Moderate the Main Effect of the Confirmation Bias

Since the paper focuses on the particularities of social media in sharing content, it steps into the question of how social media changes communication (RQ1 and RQ2). First, we see that there is a difference in perception between different account types. News media organizations are demoted to the same level as

other opinion leaders and further suffer from losing credibility and, therefore, the believability of the content they propose. Furthermore, with the sharing mechanism, it is possible to trace the originator of the content directly through messenger. Since in the paper perceived similarity is manipulated through political affiliation, here we can see how the individual level (i.e., confirmation bias) interacts with the group level (i.e., perceived similarity through affiliation to a group). Hence, the paper also provides answers for RQ3.

Paper V: Privacy is Important! Or not? – Commenting and Liking Under Confirmation Bias on Social Media[2]

Paper V focuses on the issues of privacy in conjunction with interaction. Users on social media like and comment on content, on the basis of which social media websites can derive clear preferences and behavior patterns of their users (Morales-i-Gras 2020). Therefore, these preferences and behavior patterns are personal information that can be used for various purposes, such as personalized marketing. Hence, users of social media platforms have an interest in protecting their data so that social media websites do not abuse or sell their personalized data.



Note: p-values. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.09$, ns = not significant.

Figure 15. Privacy concepts under the confirmation bias result

The paper investigates which privacy-related concepts lead to disclosure in the form of likes and comments. The concepts that were investigated are trusting beliefs in the platform (i.e., the platform will not sell or abuse the personal information) and risk beliefs (i.e., the potential negative ramifications of data abuse). Furthermore, the content presented in this case was in accordance with the beliefs of the social media users. Algorithms usually try to reinforce past information consumption. Therefore, the probability is high that content on social media confirms pre-existing beliefs (i.e., confirmation bias). The results show that under confirmation bias, only the trusting beliefs significantly influence commenting and liking, while risk beliefs become insignificant. Furthermore, the topic also plays an important role in distinguishing between liking as a form of click speech and commenting. The higher the importance, the higher the probability of commenting.

In summary, the paper elaborates on how information flow shifts from passive consumption of information to active interaction on the level of the user. Click speech in the form of liking is one new

[2] Haug, Maximilian and Gewald, Heiko (2023) Privacy is Important! Or not? – Commenting and Liking Under Confirmation Bias on Social Media. In: (Proceedings of the) 29th Americas Conference on Information Systems (AMCIS), August, 10-12, 2023, Panama City, Panama.

mechanism without the burden of articulation of one's opinion. This presentation of endorsement or opinion makes interaction more visible than it was traditionally the case. Therefore, it is also much easier to draw conclusions about individuals' preferences or attitudes. However, risks play a diminished role in this consideration, whether interaction happens or not (RQ1). Furthermore, the paper sheds light on individual-level influences on interaction, such as the personal importance of the topic in discussion, as well as the evaluation of risks and trust towards social media platforms (RQ3).

Paper VI: Supporting Opinions to Fit in: A Spiral of Silence-Theoretic Explanation for Establishing Echo Chambers and Filter Bubbles on Social Media[3]

Paper VI steps into the sphere of social media communities and the dynamics of such influencing opinion disclosure in the form of commenting. Social media communities often contain people of similar interests with a specific topic discussed within the community (Kietzmann et al. 2011). Therefore, such communities can function as echo chambers in which opinions that are contrary to the community's opinion are rejected (Cinelli et al. 2020). The paper investigates a specific mechanism to understand opinion disclosure when confronted with opinions that go against one's own but are in line with the community opinion.

The mechanism is the fear of isolation as outlined in the spiral of silence theory (Noelle-Neumann 1974). Since the fear of isolation can have trait-based and state-based characteristics, the research paper investigates their relationship and which form of the fear of isolation ultimately leads to opinion disclosure in the form of opposing or supporting the community's opinion against one's own convictions. The results show that state-based fear of isolation (i.e., the fear of isolation that occurs in a specific context and vanishes afterward) prompts community members to express agreement with the community's opinion. This happens even though the community members hold different views. These findings extend the spiral of silence theory as the theory only theorizes self-censorship and not opinion disclosure against one's own conviction. Furthermore, the paper presents topic importance as a moderating variable, which further exacerbates the relationship with topics perceived to be important. Superficial topics are seen to hold less risk of becoming a victim of isolation, while for important topics, community members are more likely to fall in line with the community opinion.

[3] Haug, Maximilian and Maier, Christian and Gewald, Heiko and Weitzel, Tim (2025) Supporting opinions to fit in: a spiral of silence-theoretic explanation for establishing echo chambers and filter bubbles on social media. Internet Research, 35 (7). pp. 30-51. ISSN 1066-2243

	Scenario 1 (immigration)		*Scenario 2 (presidential election)*		*Scenario 3 (COVID 19)*	
Independent variables	**Oppose**	**Support**	**Oppose**	**Support**	**Oppose**	**Support**
Fear of isolation (trait)	-0.057 ns	0.055 ns	-0.023 ns	0.118 ns	0.010 ns	0.000 ns
Fear of isolation (state)	-0.217 *	0.209 ns	-0.238 **	0.163 ns	-0.170 ns	0.380 ***
FoI (state) X importance	-0.219 ***	-0.019 ns	-0.126 *	-0.066 ns	-0.128 ns	-0.230 *
FoI (trait) on FoI (state)	0.694 ***		0.647 ***		0.674 ***	
FoI (trait) indirect effect	-0.150 *	0.145 ns	-0.154 *	0.106 ns	-0.114 ns	0.256 ***
Dependent variable R^2	**33.2%**	**6.6%**	**21.1%**	**6.5%**	**10.0%**	**13.5%**
FoI (state) R^2	**48.1%**		**41.8%**		**45.4%**	

Note: p-values. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.09$, ns = not significant.

Table 7. Paper VI results with Path coefficients, mediation analysis (indirect effect), R^2

The paper first elaborates on the different aspects of social media in terms of social media communities. They represent a subset of social media users with a specific focus. Hence, social media users are encouraged to engage with the content of others within such communities, highlighting their active role in the new online information ecosystem. Users are consumers and contributors with their act of commenting since other users can react and interact with the newly generated content (RQ1). Further, the paper taps into a distinct mechanism to determine opinion disclosure or self-censorship. In this specific instance, the fear of isolation was investigated, and insights on how the trait-based and state-based fear of isolation interact were provided, hence providing insights for both RQ3 and RQ4.

Paper VII: Social Media and Me: How Community Identity Influences Click Speech[4]

Building on **Paper VI**, **Paper VII** dives deeper into the relationship between social media community members and their opinion disclosure through social media mechanisms. Here, the focus lies on click speech as defined in the literature background. Furthermore, community identity as a driver is investigated in conjunction with the aforementioned fear of isolation. The aim of the paper was to contextualize the concept of community identity with the spiral of silence theory and, therefore, the mechanisms of fear of isolation. Furthermore, it investigates which of the concepts leads to click speech against one's own opinion. The spiral of silence theory suggests that self-censorship is happening when confronted with a public opinion one does not hold. In the paper, it is hypothesized that for click speech, we expect a much lower barrier to fall in line with a dominant opinion to escape potential social isolation from the community.

The results show that community identity, meaning feeling a strong connection to the community, significantly influences click speech (liking and sharing). This result shows that community identities strongly drive engagement, even if opinions are not aligned. Therefore, social media fosters constant opinion-sharing through the click speech mechanism, which further reinforces connectedness and an

[4] Haug, Maximilian and Maier, Christian and Gewald, Heiko (2024) Social Media and Me: How Community Identity Influences Click Speech. Journal of Computer Information Systems: JCIS, 64. ISSN 2380-2057

illusion of community cohesion. Furthermore, fear of isolation was contextualized in the sense that the stronger the community identity, the higher the likelihood of fearing isolation. Lastly, the influence of fear of isolation on click speech was found in two of three scenarios (COVID-19 and presidential election were significant, but the topic of immigration was not significant). At the time of the study, the first two were trending topics, while immigration was considered an ongoing issue. Therefore, hot-button issues naturally induced more engagement.

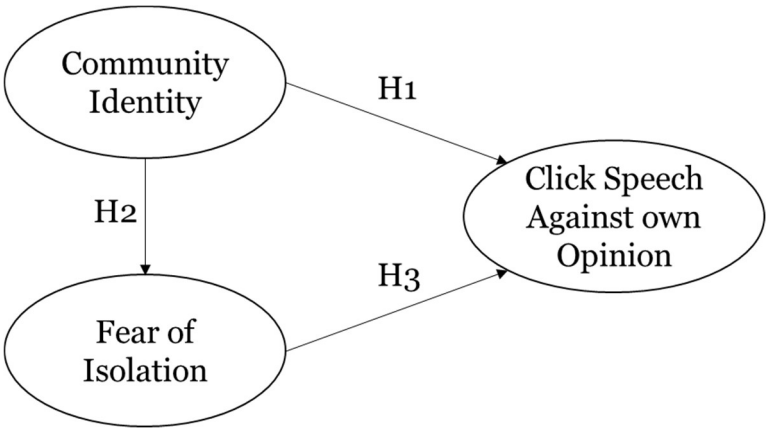


Figure 16. Research model with community identity and fear of isolation influencing click-speech

	Scenario 1 (immigration)	*Scenario 2 (presidential election)*	*Scenario 3 (COVID 19)*
Influence on Click Speech	**Click Speech**	**Click Speech**	**Click Speech**
Community Identity (H1)	0.276 ***	0.208 **	0.242 **
Fear of Isolation (H3)	0.147 ns	0.222 **	0.271 **
Influence on Fear of Isolation	**Fear of Isolation**	**Fear of Isolation**	**Fear of Isolation**
Community identity (H2)	0.251 **	0.263 **	0.195 *
Dependent Variable R^2	**11.8%**	**11.7%**	**15.8%**
Fear of Isolation (state) R^2	**6.3%**	**6.9%**	**3.8%**

Note: p-values. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.09$, ns = not significant.

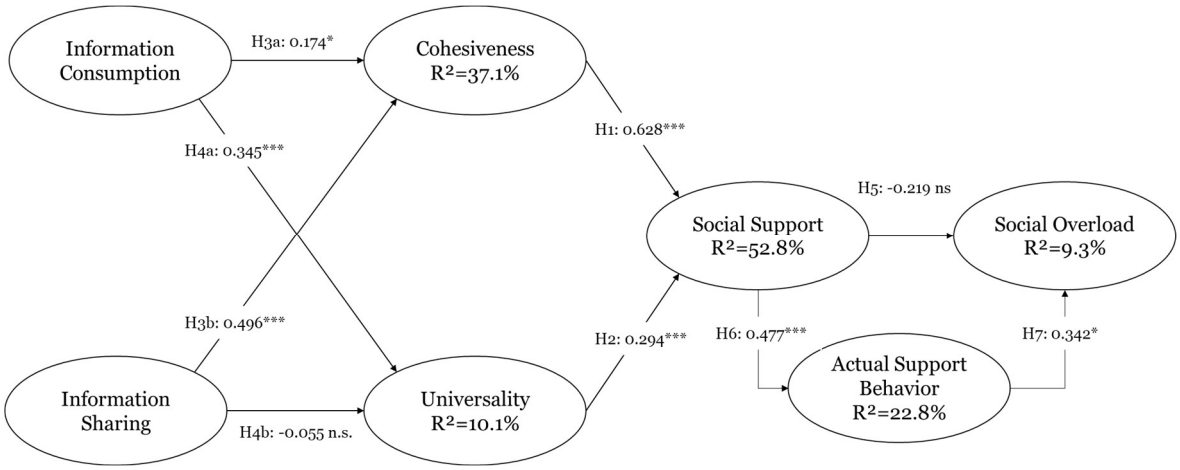
Table 8. Results for community identity and fear of isolation on click speech

In summary, the paper provides insights into how social media fosters cohesive communities, which in turn cultivate opinions that are not necessarily aligned with the true beliefs of its members. The stronger ties among the users as audience therefore show a stark contrast in communication change, since such cohesive communities were not possible across multiple topics before the rise of social media (RQ1). For the mechanisms of group-level drivers for engagement within social media communities (RQ3, RQ4), the paper identifies community identity as a major driver, even if members do not align with the presented opinions. Protecting the community identity and the fear of not becoming isolated leads to

interaction against one's own opinion. Lastly, the paper sheds light on click speech as a distinct mechanism of opinion disclosure, which provides further answers for RQ5 in conjunction with Paper VI, which focuses solely on commenting.

Paper VIII: Balancing Taking and Giving: Contextualization of Social Support and Social Overload in Online Mental Health Communities[5]

In **Paper VIII**, the focus lies on how interaction may be inhibited by extensive participation in social media communities. In that regard, the concepts of social support and social overload were investigated. Specifically, their relationship occurs when users interact in online mental health communities as a context. As mentioned in the literature background, community members receive social support, but also can give social support. Research, however, showed that in such a community, social overload, the feeling of giving too much, can occur.



Note: p-values. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.09$, ns = not significant.

Figure 17. Determinants of social overload with social support and antecedents

The results show that information consumption within such a social media community influences cohesiveness (i.e., staying in touch with other community members) and universality (i.e., the recognition of experiences of other community members). Furthermore, information sharing influences cohesiveness. Both concepts, cohesiveness and universality, significantly influence social support. This means that with a strong attachment to the community and its members, community members have the feeling of receiving social support. Due to the reciprocal nature of interaction, a relationship between receiving social support and actual social support behavior was hypothesized. This relationship was found to be significant. While receiving social support does not directly influence social overload (meaning the reason for receiving social support might induce social overload, since members may feel

[5] Haug, Maximilian and Tikil, Alisha and Gewald, Heiko (2025) Balancing Taking and Giving: Contextualization of Social Support and Social Overload in Online Mental Health Communities. In (Proceedings of the) 33rd European Conference on Information Systems (ECIS), June, 12-18, 2025, Amman, Jordan

the obligation to give something back), only the actual supporting behavior was found to induce social overload.

Paper VIII presents various insights that are relevant to answering the research questions. First, social media communities provide benefits for the interaction of their users. Social support is one of the main benefits for social media users. The study shows how social support emerges in these communities, which is a key factor in user interaction. Cohesiveness and universality act as influences for social support and therefore provide insights into factors that drive social media community interaction. On the other hand, giving social support and the associated social overload act as inhibitors for interaction (RQ4). Furthermore, the paper taps into the negative ramifications of social media use within social media communities. Namely, social overload, as the dark side of excessive interaction, induces stress in users (RQ6).

Paper IX: Content Creators on Instagram – How Users Cope with Stress on Social Media[6]

Paper IX fully steps into the ramifications of social media interaction in the context of unintended consequences. Hence, we observe the dark side of social media interaction on an individual level. The basis of the paper is coping theory (Folkman et al. 1986), and the goal was to understand what factors cause stress in the context of social media interaction, independent from communities. Even though literature investigates social media as a coping tool, research lacks insights into the specific strategies users enact after having a negative interaction. Furthermore, the paper elaborates on what a negative interaction possibly can contain.

As outlined in the methods section, the paper is based on a qualitative dataset that investigates the origins of stress. The data suggests that a lack of likes on one's own content, especially, induces stress in users. Before posting content, users have a certain anchor of how many likes they may expect they receive. If the actual likes undercut this anchor drastically, users perceive this as a negative interaction. Furthermore, negative user comments were also mentioned as the origins of negative interaction. However, insufficient likes were vastly more prominent. Further, coping strategies differed based on the origin of the negative interaction. Some strategies were only enacted due to insufficient likes, while others were only enacted due to negative comments. The following table provides an overview of the different coping strategies and categorizes them based on whether they address the root problem (problem-focused) or manage emotions (emotion-focused).

[6] Haug, Maximilian and Reiter, Julia and Gewald, Heiko (2024) Content creators on Instagram - How users cope with stress on social media. Telematics and Informatics Reports, 13. ISSN 2772-5030

Emotion-Focused		Problem-Focused	
Coping Strategy	*Stressor*	*Coping Strategy*	*Stressor*
Distraction	Likes, Comments		
Blocking	Comments	Changing posting behaviour	Likes
Content deletion	Likes, Comments	Mindset change	Likes, Comments
Action out of spite	Likes	Breaking habits	Likes
Defense	Comments	Deletion of app	Likes, Comments
Self-sabotage	Likes		

Effort ↓ | ↑ Facilitates Content

Table 9. Coping strategies with associated stressors and effort (emotion-focused) and content facilitation (problem-focused)

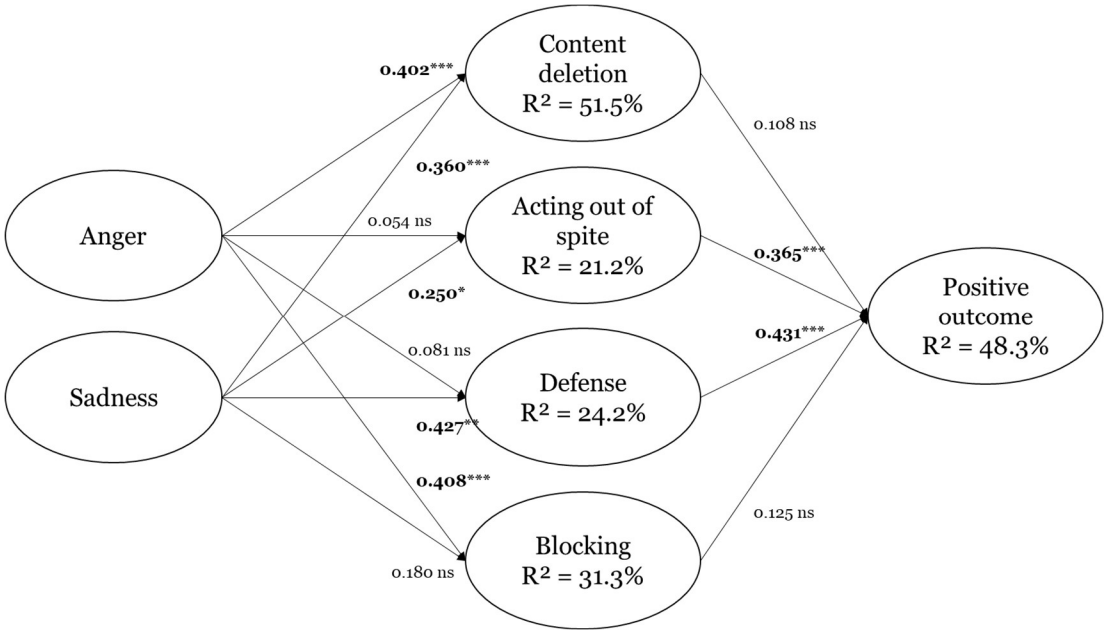
In summary, the paper investigates origins of stress arising within users interacting on social media. With insufficient liking as one major reason and negative comments, the paper provides answers for RQ6. Furthermore, the second major aspect of the paper was the identification of possible coping strategies in order to overcome the stress induced by likes and comments. With the distinction of emotion-focused and problem-focused coping strategies, the paper provides answers for RQ7.

Paper X: Anger and Sadness – Coping Strategies to Manage Negative User Interaction on Instagram[7]

Finally, **Paper X** builds on the results of **Paper IX** and quantitatively investigates whether the mentioned coping strategies actually show efficacy in their effectiveness. The paper takes an emotion-based approach by elaborating on which negative emotions, namely anger and sadness, arise due to a negative user interaction on Instagram. As outlined before, the negative interaction is operationalized as insufficient likes and a negative comment. As coping strategies specifically four were chosen as they can directly be enacted on social media: Content deletion, as the deletion of the post associated with the interaction; acting out of spite, meaning continue to post the same content without consideration of the interaction; defense, meaning arguing and justifying the content in the comment section; and blocking as in blocking users with that post negative comments.

The results show that such negative user interaction sparks high and low arousal negative emotions such as anger and sadness. Furthermore, anger showed to facilitate rather destructive coping strategies such as content deletion and blocking. These do not further increase content creation. Sadness, on the other hand, leads to acting out of spite and defense, which anger did not facilitate. Hence, negative emotion also leads to further content creation on social media. Lastly, the destructive coping mechanisms show no significant influence on positive outcomes, as reducing stress. The strategies that facilitate content creation show efficacy in reducing stress.

[7] Haug, Maximilian and Reiter, Julia and Gewald, Heiko (2024) Anger and Sadness — Coping Strategies to Manage Negative User Interaction on Instagram. In: (Proceedings of the) 32nd European Conference on Information Systems (ECIS), June, 13-19, 2024, Paphos, Cyprus



Note: p-values. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.09$, ns = not significant.

Figure 18. Structural equation modelling results, Significance key: *** p < 0.001; ** p < 0.01; * p < 0.05

Paper X dives into the efficacy of coping strategies that can be enacted directly on social media platforms. It provides further context to the ramifications of negative user interaction due to insufficient likes and negative comments (RQ6). Such ramifications are the emergence of negative emotions such as anger and sadness. Therefore, the paper establishes a link between negative user interaction and coping strategies through the concept of emotions. Further, the paper quantitatively tests the previously identified coping strategies and ultimately shows which ones actually show efficacy as coping strategies (RQ7).

7 Contributions and Implications

7.1 Contribution to Theory

In the following, the contributions to theory are outlined based on the individual research questions.

7.1.1 Changes of Communication and Empowerment through Social Media (RQ1)

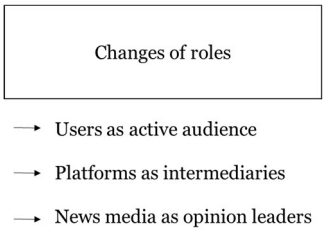
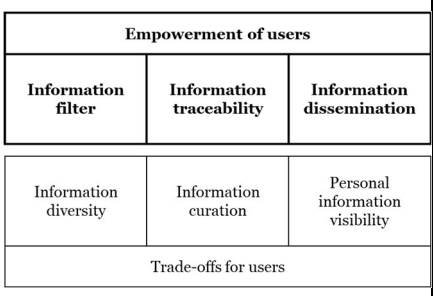
When we consider the communication literature, we find a multitude of different models that explain how communication between individuals can work. As outlined in the literature background, Westley and MacLean Jr (1957) consolidated the "jungle of communication models" in their model to make sense of the literature. With the emergence of social media platforms, IS literature tapped into their central role in the online information ecosystem as new media (Dennis and Kinney 1998). Platforms are characterized by their highly synchronous communication patterns (Dennis et al. 2008), which comprise actors and their different relationships (Jin et al. 2019). However, literature lacks a conceptualization of the changes in communication through social media and the ramifications on a user level.

The changes in communication occur mainly through changes in roles and patterns as outlined by **Papers I** and **III**. There are three major changes. First, the audience changed from a traditionally passive consumption-oriented mass (Westley and MacLean Jr 1957) to an active user base that interacts with content and other actors on the platform. Second, social media platforms take the role of the intermediary as the new gatekeeper of information through manual and automated content moderation. Third, news media, formerly the gatekeepers of information, lose power and assume the role of opinion leaders among traditional opinion leaders, such as politicians or activists. Hence, this dissertation conceptualizes these new roles in the online information ecosystem. As this work focuses on users on social media platforms, various ramifications stem from these changes in roles.

This dissertation contributes to a conceptualization of the empowerment and trade-offs on the side of social media users as the audience by combining results from the presented papers. **Paper I** specifically investigates the changes of roles within social media and the empowerment of the user. With social media platforms and the possibility to follow people or organizations of interest, users can actively control what kind of content they want to interact with. **Papers VI** and **VII** provide further context for consciously chosen information environments. By stepping into social media communities, we find that users of similar interests gather in online spaces to interact with each other, actively shaping their information diet. Furthermore, **Paper I** elaborates on social media algorithms as they passively influence their information diet through what users predominantly consume and interact with. **Paper IV** highlights the possibility of tracing information to its origin, since social media provides information about who shared which information. Lastly, a major change is the shift from the passive to the active audience, since users are now part of the dissemination process through social media inherent mechanisms such as liking, sharing, and commenting.

With empowerment, there are certain trade-offs that facilitate unintended consequences and responsibilities on the user's side. **Paper I**, with its focus on fake news, highlights that the power to determine the information diet poses a high risk of a reduction in information diversity. Users follow people or organizations or participate in communities that show a high overlap in their own convictions. Social media algorithms elevate this tendency as they are tailored to maximize interaction, which is based mainly on confirmation bias. Hence, social media platforms enact an algorithmically enforced confirmation bias. With the aforementioned changes in news media in the online information ecosystem, there is a stronger emphasis on information curation on the side of the user. It becomes increasingly difficult to evaluate the credibility of actors and the veracity of information. **Paper IV** shows how the different perceptions of accounts influence the believability of content. **Paper II** shows that such evaluations are heavily in favor of heuristics, such as the perception of the profile that shares content. Perceived similarity and confirmation bias are the main drivers in this regard. Lastly, as information is easily traceable, users become more transparent, as outlined in **Paper V**. Such transparency infringes on the privacy of users. However, users also enact heuristics, such as trusting the platform not to abuse their data and discounting privacy risks in the future.

In summary, we see that social media provides empowerment for users, but also trade-offs. The results show that there is high potential for users to tailor the online information ecosystem in their favor. However, the trade-offs show that users do not adequately account for their empowerment, as their behavior mainly rely on heuristics that are based on the perception of profiles that share content they are interested in. Hence, social media users need to focus on the trade-offs and responsibility to harness their newly gained empowerment.

<table>
<tr><th></th><th></th><th>This dissertation ...</th></tr>
<tr><td>Social media changes the role of communication.</td><td> 
</td><td>Reveals the changing nature of communication roles:
Shift of the users from a passive to active audience through liking, sharing, and commenting
Social media platforms are new gatekeepers of information through algorithms and active moderation
Loss of power of news media as they now assume the role of opinion leaders</td></tr>
<tr><td>Social media empowers its users.</td><td> 
</td><td>Identifies the empowerment of users:
Users can tailor their information diet through information filters (active) or through what they dominantly consume (passive)
Users can easily trace information within the online information ecosystem due to information containing who is sharing or liking which content
Users disseminate information on their own through the liking, sharing, and commenting mechanisms</td></tr>
</table>

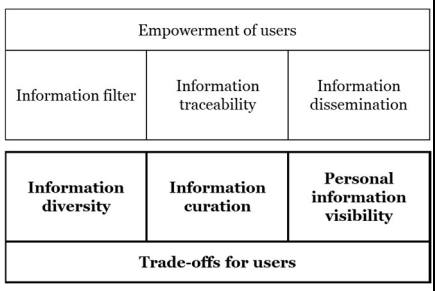
With empowerment, there are trade-offs and responsibility for users.		• Shows self-made echo-chambers through the lack of information diversity based on algorithmically enforced confirmation bias and social media communities • Reveals the shifting responsibility of information curation from news media to the users, with emphasis on heuristics to determine credibility and veracity • Shows that users also rely on heuristics to discount privacy issues on social media by relying on the trustworthiness of the social media platform

Table 10. Contributions concerning changes of communication through social media

7.1.2 Communication through Sharing with Originators and Messenger (RQ2)

Literature shows that presentation format has a strong influence on how content is evaluated by users (Kim and Dennis 2019). Therefore, we already know how content specifics influence concepts like believability. Beyond that, as established in the previous section, users heavily rely on heuristics to evaluate content on social media. This is dominantly based on the perception of different account types. Therefore, to understand the ramifications of sharing, this dissertation first sets the basis of conceptualizing the three main account types with which users interact when content is shared. **Paper I** identifies different account types from the literature, while **Papers II** and **III** test the perception of such account types. First, news media organizations, formerly gatekeepers, are still perceived differently from traditional opinion leaders. They are characterized by their main agenda: news reporting, and thereby make the best effort to appear impartial and credible. Second, traditional opinion leaders, as outlined by literature (Westley and MacLean Jr 1957), are advocates for their cause and thereby consist of political individuals and parties, organizations, activists, and influencers. Lastly, as a main new and relevant account type, users step into the information flow process as active disseminators of information. These consist of accounts that suggest using social media to inform themselves and to interact with others. In summary, these three account types (i.e., news media organizations, traditional opinion leaders, users) have different main foci that distinguish them based on which agenda they follow on social media.

Based on the different account types, social media provides a multitude of possibilities for users to interact with content through the sharing mechanism. **Papers II** and **III** highlight the differences between immediate and mediate communication patterns through different account types. Adding to the literature on presentation format (Kim and Dennis 2019), the papers show that the perception of believability of content is additionally based on the originator's account and the messenger of such content. In particular, there are two effects of mediated communication. First, there is a messenger effect that shows that the perception of the account type that shares content significantly alters the perception based on perceived similarity. Second, if in addition the originator account is highly dissimilar, the believability even further increases, hinting at a mixed-actor-amplification effect.

In summary, this dissertation contributes by conceptualizing the main account types on social media based on how users perceive them. Based on the account types, different perceptions of content occur through immediate and mediated communication patterns.

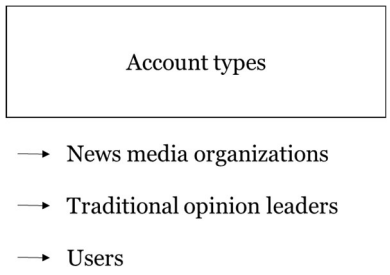
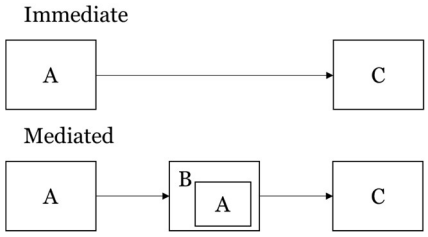
		This dissertation …
Emergence and Conceptualization of account types that share content		Conceptualizes three different main account types on social media based on their intentions • News media organizations, as former gatekeepers, are now building on their credibility – enact news reporting • Traditional opinion leaders such as politicians/parties, activists(organizations), or influencers – driving their own (political/business) agenda • Users as private individuals with the main goal to interact with other actors to receive information and build relationships – connection and interaction
Differentiation between immediate and mediated communication		Highlights the difference between immediate and mediated communication • Perception of content changes based on immediate or mediated interaction • Messenger effect: the messenger (B) dominantly determines the believability of content based on perceived similarity • Mixed-actor-amplification-effect: similar messenger (B) sharing from dissimilar originator (A) shows the highest effect in believability

Table 11. Contributions concerning communication through sharing

7.1.3 Interaction through confirmation bias, perceived similarity, and fear of isolation (RQ3)

Literature shows that confirmation bias is a strong driver for interaction (Kim and Dennis 2019; Moravec et al. 2019). Such confirmation bias hinges strongly on the content of information presented on social media. This dissertation contributes to this literature stream by extending our understanding of confirmation bias by investigating its interaction with perceived similarity to the actor that shares content. Hence, this work elaborates on the fact that interaction happens on the basis of content and the actor sharing the content. Specifically, perceived similarity directly influences the perception of content, as outlined in **Paper III**, but it also moderates the relationship of confirmation bias on the perception (believability), as outlined in **Paper IV**. These findings are in line with **Paper II**, which suggests that users on social media use the perception of the individual account types as a heuristic to determine the credibility of content. Perceived similarity acts as an exacerbator and dampener of believability. Low similarity reduces believability, while high similarity increases it. However, in the case of mediated communication, low similarity in the context of the originator, while high similarity for the messenger, shows the highest influence on believability. Hence, this work contributes by proposing a configuration of similarities through the communication process that increases or decreases believability.

As a further contribution, **Paper VI** investigates different forms of the fear of isolation mechanism from the spiral of silence theory (Noelle-Neumann 1974). As mentioned in the literature background, the mechanism of the spiral of silence can have trait and state-based characteristics. However, literature has not yet investigated how these two concepts interact with each other. **Paper VI** reveals that the trait-based part of fear isolation alone does not determine interaction in social media communities. State-based fear of isolation, however, mediates the trait-based concept. Hence, there is a trait-state mediation for these two concepts. Furthermore, the literature has not yet provided a clear measurement tool for the state-based fear of isolation. Hence, new items were developed in order to clearly differentiate between trait-based and state-based fear of isolation.

Since the state-based fear of isolation was shown to dominantly influence interaction, its relationship with the established influence of confirmation bias was tested. Here, **Papers VI** and **VII** showed that the mechanism of the state-based fear of isolation interacts in a way with confirmation bias, as it crosses out its influence. This finding adds to research as the literature dominantly focused on the topic, the account who shared the information, and how information was presented in general (Kim and Dennis 2019; Kim et al. 2019). Now, under specific circumstances, such as social media communities, we see that social-related influences heavily infringe on the influence of confirmation bias to the point that confirmation bias is completely crossed out. The resulting interaction extends the spiral of silence theory as its main prediction was self-censorship when confronted with threatening isolation. Here, it is shown that active opinion disclosure against one's own belief system can be the outcome of fear of isolation in the context of the social media community.

These findings open up the discussion for a hierarchy of influences on interaction on social media. In the public space of social media, confirmation bias and perceived similarity seem to be dominant drivers for interaction. The dissertation revealed that the stronger the context shifts towards users interacting in a setting that is associated with a community in which social punishment can be anticipated, the higher the likelihood that the concept of confirmation bias is overruled by the social pressure. Specifically, the fear of becoming isolated weighs higher than one's own convictions. As a result, social media communities can contain individuals who will support opinions they do not even hold. Such fear of isolation is exacerbated by a person’s identity associated with the community. The combined findings suggest, therefore, a hierarchy of influences, in which individual perceptions of content (confirmation bias) and account (perceived similarity) set the baseline, but ultimately, the social pressure in the form of potential punishment if individuals do not fall in line supersede this baseline.

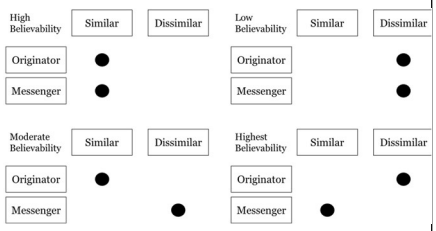
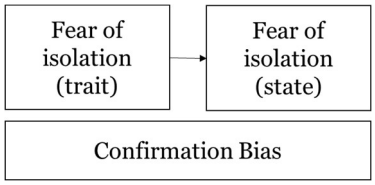
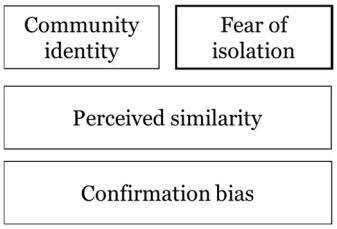
		This dissertation …
Perception of content through confirmation bias and perceived similarity		• Shows that the perception of content on social media is based on content plus the actor • Reveals perceived similarity as an exacerbator or dampener of confirmation bias • Highlights the interaction term of originator X messenger and proposes a configuration of similarities in the communication process
Contextualization of fear of isolation with confirmation bias		• Contextualizes trait-based and state-based fear of isolation in a trait-state-mediation • Develops a measurement for state-based fear of isolation • Highlights the interaction of confirmation bias and fear of isolation on interaction on social media
Hierarchy of influences		• Reveals that interaction is heavily context-specific (social context) • Shows that individual differences set the baseline, which is then influenced by group-based influences • Shows how fear of isolation and community identity supersede confirmation bias and perceived similarity

Table 12. Contributions concerning interaction through individual and group-based influences

7.1.4 Major Social Influences of Interaction in Social Media Communities (RQ4)

Social media platforms provide the possibility to create communities as distinct places of communication on social media (Karahanna et al. 2018; Kietzmann et al. 2011). Research on what drives interaction in such groups is scarce since most literature focuses on the public aspect of social media as a place to connect with millions of other people (e.g., Kim and Dennis 2019). Social media communities naturally contain a more socially interconnected space in which their members may hold strong social ties. Hence, this dissertation investigates major drivers of interaction within social media communities, with the focus on drivers that show a social aspect. **Papers VI** and **VII** identify fear of isolation as a strong influence, as it shows the capability for interaction against one's own beliefs. **Paper VII** furthermore also shows that community identity has a direct influence on interaction. **Paper VIII** highlights how receiving social support can trigger the motivation to interact with the community to "give something back". The paper further shows how social support forms in such communities through universality and cohesiveness.

Furthermore, this dissertation contextualizes identity theory within the spiral of silence theory as outlined in **Paper VII**. Specifically, the mechanism of fear of isolation (state) and its relationship with community identity is elaborated on. The research shows that community identity has a direct influence on interaction, but also is mediated by fear of isolation. This is important to understand on what basis

fear of isolation emerges within social media communities. Therefore, this work highlights community identity as an antecedent of fear of isolation.

Lastly, the spiral of silence was postulated to lead to self-censorship. The proposed research papers show that in social media communities, the mechanism of fear of isolation does not lead to self-censorship but to opinion disclosure against one's own beliefs. However, literature shows further concepts that inhibit interaction. When we look at the dark side of social media, we find social overload as an influence that reduces interaction. This dissertation contributes by disentangling the nuances of how social overload occurs (**Paper VIII**). Specifically, it shows that not the mere receiving of social support strains community participants, but at the point they give social support through interaction, social overload manifests, and in return inhibits further interaction.

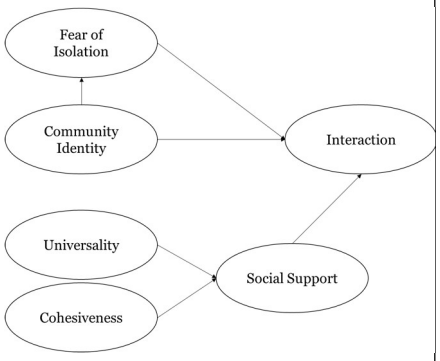
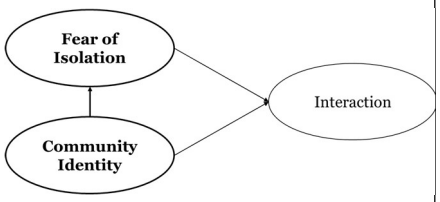
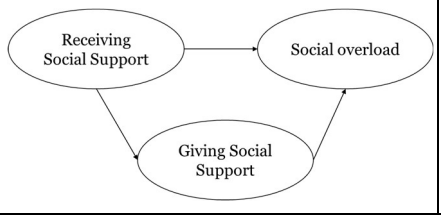
		This dissertation …
Identification of major socially associated drivers		• Identifies major influences on interaction within the context of social media communities • Shows community identity and fear of isolation as major drivers for interaction • Highlights the role and influences of social support that lead to interaction • Shows that universality and cohesiveness influence social support
Identity and fear of isolation		• Contextualizes community identity within the spiral of silence theory • Shows how community identity contributes to opinion disclosure against one's own beliefs • Shows how fear of isolation mediates community identity
Inhibitors of interaction		• Highlights social overload as an inhibitor of interaction within social media communities • Contextualizes social support and social overload through actual behavior in giving social support

Table 13. Contributions concerning influences of interaction in social media communities

7.1.5 Interaction Through Click Speech and Commenting (RQ5)

Social media platforms provide distinct forms of interaction behavior, which mainly come down to click speech and commenting (Robbins 2014; Sklan 2013). This dissertation highlights the differences between the two forms of interaction with content in the context of social media communities. This is

important since click speech metrics are a key characteristic of content to determine its popularity and reach (Van Aelst et al. 2017).

First, this dissertation highlights the distinction between the two forms of interaction on social media and specifically within social media communities. As outlined in the literature, click-speech is an easy-to-enact form of interaction, which this dissertation confirms and expands on in the context of social media communities. The results of **Papers VI** and **VII** combined show that within social media communities, click speech is an easy-to-enact form of interaction with context in the context of avoiding social sanctions. Specifically, participants of social media communities see liking, through a simple click, as an easy way to signal they are still part of the community, even though the participants may not endorse the proposed content. The papers show that this opinion disclosure against one's own beliefs is more prevalent with click speech than with commenting. This indicates a lower cognitive barrier for click-speech, since it is only one click and does not hold the same consideration of the content that may be associated with commenting. Lastly, this dissertation contextualizes click-speech as a form of opinion disclosure within the spiral of silence theory. This is important to understand, since the spiral of silence theory does not elaborate on different forms of opinion disclosure, as it is now the reality on social media platforms. Hence, there is more nuance to consider when investigating interaction on social media.

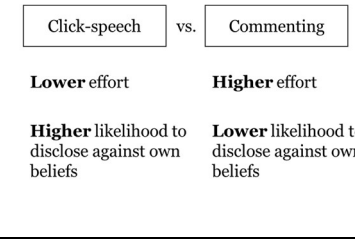
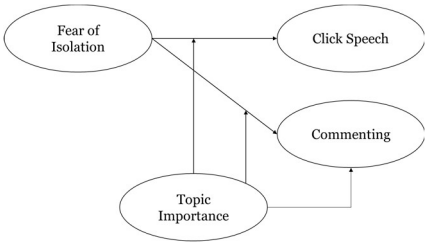
		This dissertation …
Click speech vs commenting		• Highlights click speech as an easy-to-enact form of interaction with lower barriers than commenting • Shows that within social media communities, participants are more likely to betray their own opinions in a social setting in the form of click-speech than in the form of commenting • Contextualizes click speech within the spiral of silence theory.
Topic importance as a moderating and influencing variable		• Contextualizes topic importance as a content-specific characteristic within click speech and commenting behavior within the spiral of silence theory • Highlights topic importance as a moderating variable that exacerbates fear of isolation on opinion disclosure against one's own beliefs under high topic importance • Highlights topic importance as a direct influence on commenting in contrast to click-speech

Table 14. Contributions concerning interaction through click speech vs. commenting

Second, the interaction with content against one's own beliefs is moderated by the topic. **Papers VI** and **VII** show that different topics lead to differences in interaction in terms of click-speech and commenting. The results indicate that the support of opinions one does not hold is an interaction term between the topic in discussion and the social component, in this case, the fear of isolation. **Paper V** further provides context on the direct effect of topic importance, since the paper showed that only commenting is significantly influenced by topic importance. This underlines the notion that there is less

elaboration on the topic when it comes to click-speech, and such interaction behavior is mainly used to connect socially in the context of social media communities.

7.1.6 Negative Impacts of Social Media Interaction (RQ6)

This dissertation goes beyond the interaction on social media and highlights the ramifications of negative interaction as a part of the research stream of the dark side of technology (e.g., Beyens et al. 2016; Maier et al. 2015). This work elaborates on two perspectives. The first perspective focuses on stress emerging from being in a reciprocal environment, which is the case within social communities. Community members receive and give social support. **Paper VIII** shows that with the giving of social support, social overload can emerge. Such social overload then stands in stark contrast to what community participants actually expect from interaction within the community. This mismatch fits the definition of stress in terms of Folkman et al. (1986). Hence, social overload acts as an origin of stress in such community settings. Furthermore, the paper shows that it is not the perception of being expected to "give something back" but the actual supporting behavior that triggers social overload. This is important to understand since this shows that the members who care and interact are the first to experience such unintended ramifications.

The second perspective focuses on specific metrics (i.e., likes) associated with content. **Papers IX** and **X** outline likes and comments as salient types of interaction that constitute either a positive or negative interaction. Shares were not mentioned as an indicator of negative user interaction. Both papers highlight how social media users who post content show a certain threshold of likes, which they consider mandatory, in order for them to conclude that the interaction with other users with their content is positive. If such a threshold is not met, the interaction is deemed as negative, and users become stressed as their expectation mismatches with reality. Likes in that regard are seen as the first and most important metric that constitutes how interaction on social media is received by content creators. Comments also play a role, but are not seen as a mandatory occurrence. The absence of comments does not constitute a negative interaction. However, negative comments influence perception and funnel into what is seen as a negative user interaction.

		This dissertation …
Social overload in social media communities		• Highlights the role of social overload within social media communities as the origin of stress • Shows how social overload emerges through supporting behavior
Likes and comments	Likes → Mandatory to reach threshold → Heavily associated with negative interaction Comments → Not expected → Negative comments constitute negative interaction	• Shows how social media users have expectations for a like threshold for their content • Highlights the main driver for the perception of what constitutes a negative interaction on social media • Shows that users do not expect positive comments, but see negative comments as the origin of negative user interaction

Table 15. Contributions concerning reasons for stress in social media Interaction

7.1.7 Coping Strategies with Negative Interactions on Social Media (RQ7)

Literature suggests that social media can act as a coping tool for stress (Van Ingen et al. 2016). As social media enables users to connect and exchange information (Karahanna et al. 2018), it is not surprising that the platforms can be used as a resource to cope with stress. However, as outlined before, the interaction on the platform can also be a source of stress. Hence, users develop strategies to cope with stress originating from social media platforms. This dissertation contributes to research by shedding light on distinct coping strategies arising from different origins of stress (insufficient likes and negative comments). **Paper IX** explicitly outlines emotion-focused and problem-focused coping strategies, which social media users enact on the platform but also beyond the platform in offline settings.

Furthermore, **Paper X** shows how stress manifests in form of negative emotions. Hence, this work provides further details on what constitutes stress on social media when faced with negative user interaction. Namely, high- and low-arousal emotions, such as anger and sadness emerge from the misfit between expectation and reality.

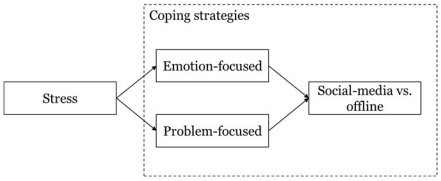
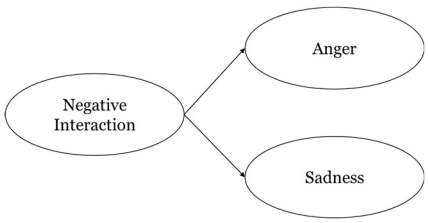
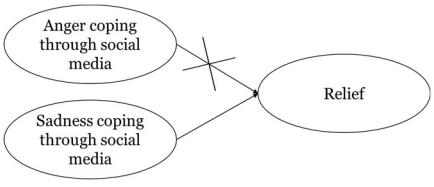
	This dissertation …
Coping strategies from negative user interaction		• Identifies different emotion-focused and problem-focused coping strategies for negative user interaction
Anger and sadness as manifestations of negative user interaction		• Sheds light on emotional responses after negative user interaction • Contextualizes low- and high-arousal negative emotions in the transactional model of stress and coping
Efficacy of coping strategies		• Identifies the efficacy of coping strategies enacted on social media • Highlights how social media-related coping strategies cannot deal with high-arousal negative emotions such as anger • Highlights coping strategies that facilitate more interaction, reduce low-arousal negative emotions such as sadness

Table 16. Contributions concerning coping with stress on social media

In a last step, **Paper X** tests the coping strategies found in **Paper IX** and shows which ones actually lead to a feeling of relief after enacting them. The dissertation contributes by showing that anger-related coping strategies are rather destructive and do not lead to positive mental outcomes. Sadness-related coping strategies lead to coping that facilitates more content and interaction, by engaging with users, or

by continuing to post content in the hope that future interaction will be positive. This is important to understand, since this contribution shows that also negative emotions, such as sadness, can facilitate further user interaction on social media.

7.1.8 Synthesis of Theoretical Contributions

This section provides an overview of the theoretical contributions across the individual chapters. As outlined in the following figure, the main contribution of this work is to provide a cascading framework that lays out changes of communication, resulting interaction, and ramifications of interaction. (1) First, in chapter 1, major changes of communication are identified with the main point of empowerment, but also new responsibilities of ordinary social media users. (2) These lead to users being now part of the dissemination process of information, and through these new forms of interaction (i.e., sharing, click speech, commenting). (3) As users switch their role from passive to active, new outcomes, in the form of unintended consequences, emerge on the individual level, mainly based on likes or the feeling of becoming socially overloaded, which need to be addressed by each social media user.

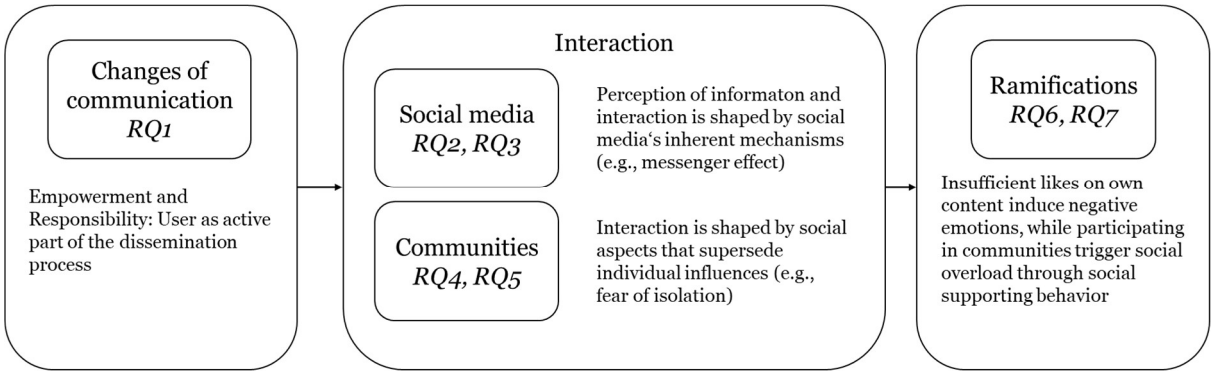


Figure 19. Synthesis of theoretical contributions

7.2 Practical Implications

The findings of this dissertation also have practical implications. The practical implications are structured into three parts. First, implications due to the change of roles and patterns in the online information ecosystem. Second, the focus is on the users and the practical implications on their side. Lastly, practical implications in reference to the unintended consequences of social media interaction are presented.

7.2.1 New Patterns and New Roles with Different Responsibilities

Nowadays, social media platforms are intermediaries that connect people on the internet. Hence, news media organizations lost their unique status as gatekeepers, having the privilege to access a huge audience. As news media organizations find their new place as opinion leaders, among the traditional opinion leaders, there is a perceived loss of credibility from the perspective of users (**Paper I**). However, we see that opinion leaders who are not associated with news media gained more credibility than before,

while news media lost a substantial amount of their credibility. As the findings show, it is important who is sharing information (**Papers II** and **III**). Hence, for news media organizations, a strategy going forward is to encourage their audience to share their content further. This increases reach, but also believability. Furthermore, the mixed-actor-amplification-effect signals that it is also important to interact with dissimilar actors in cases where opinions align. Dismissing such interaction is a lost opportunity to rebuild credibility.

For intermediaries, practical implications and suggestions face a certain dilemma (**Paper I**). Social media platforms have an inherent motivation to increase interaction to enact their business model (Tufekci 2014; Zuckerberg 2018). Hence, content that amplifies user interaction is naturally boosted. In this equation, the type of content is not relevant until it reaches content that is prohibited by law. Currently, we do not see that platforms show their own interest in changing their algorithms to create a healthier online information ecosystem. There may be a chance that in the future, we will see platforms that exacerbate extreme content also facilitate extreme user bases to the point where moderate users exile themselves, which in return would halt the growth of the platform. Until such a case manifests, further possibilities are for regulation through governmental intervention. Such intervention, however, also has its downside, as on what value basis will platforms be moderated? For example, the USA and Europe have different understandings of what speech ought to be protected and what is not. Therefore, such regulation could be unwieldy.

As users became empowered (**Papers I, II**, and **III**) to be also disseminators of information through their interaction with content, we see a shift in their responsibility. Information curation has become an integral activity of social media users, and it is also a skill to verify information. Paradoxically, users rely much more on heuristics that involve the credibility of the actor, even though the focus should lie on the content and verification behavior. These two vulnerabilities (more responsibility and still clinging to heuristics) exacerbate unintended consequences such as Fake News or the shift to echo chambers. Also, in this regard, it is very difficult to provide one solution. However, small solutions can shift the needle in the right direction, such as community-based verification of content, which shifts the focus away from the actor towards the content itself, as the community is too amorphous an entity. Furthermore, an inoculation for false information can also be implanted, meaning users are being exposed to small amounts of false information to understand the nature of such content and possible strategies malicious actors use. This experience, in return, will have a protective effect going forward when engaging with false information.

7.2.2 Social Ties Dictate Information Flow on Social Media

This dissertation highlights the force of social ties within social media interaction and the ramifications for the online information ecosystem from an opinion-forming point of view. Participants of social media communities (**Papers VI** to **VIII**) need to understand that their interaction behavior significantly influences the perception of opinions in general. Users are now part of the information flow and hence a substantial part of what we would call “the media” (**Paper I**). Hence, as mentioned beforehand, users need to embrace their new role within this ecosystem and consciously engage within

it, actively pushing for civil discourse. Community-checked information is the first step in involving communities more closely in the dissemination and curation of information. However, further steps are needed to create awareness of how important user interaction is in shaping the online information ecosystem.

As communities on social media are distinct spaces, moderation can also happen within such spaces on a community level. Here, moderation teams also need to be aware of how they shape discourse within such communities. Setting rules that allow or encourage diversity of opinions may break the fear of isolation, as community participants see that not only is one single opinion held by the community. However, this research also shows that dominant community opinions may only be a mere perception, and communities have vastly diverse opinions. Such opinions are dampened due to the strength of the social ties. When community members become aware that such diverse opinions actually exist, they may also be more willing to come forward in discussing them. Communities could implement opinion barometers, which are completely anonymous, to show how the community stands on certain positions. This way, diversity of opinions is made salient, and the effects mentioned in this work (i.e., opinions are being supported that one is not holding) can be reduced.

7.2.3 Unintended Consequences through the change of the Online Information Ecosystem

As communication changes due to the online information ecosystem, we are confronted with unintended consequences of that change (**Papers VIII** to **X**). Such consequences can be observed on a societal level and on an individual level.

On a societal level, we see phenomena such as Fake News, mis- and disinformation as a result of the changes in communication. Since news media now compete with other opinion leaders and there is a general reduction in trust, information on social media has become less reliable. Furthermore, users engage with such information based on heuristics (**Paper II**), which accelerates the issue, since information is generally not checked, but instead further disseminated. To counteract phenomena such as Fake News, various countermeasures can be implemented, as there is no single solution. These activities may incorporate fact checks, community notes, and information curation on the platform side. On the user side, there is a dire need to foster media literacy and an awareness of the new responsibilities within the communication process. Also, governments can intervene and demand more transparency on how platforms curate information and ultimately set the legal framework for what content shall be allowed. This endeavor is a delicate one, as one always needs to consider the tradeoffs between counteracting malicious actors and phenomena and the ability to express oneself freely.

Lastly, on the individual level, there is a major change in how the online information ecosystem can harm users. First, this dissertation showed that with social media platforms, there is an increased velocity and quantity of information, which may lead to information overload. This is also the reason why users tend to use heuristics to make sense of the information flood. Second, social media communities can impose strong social pressure on individuals, which leads them to take stances they do not even hold. Third, interaction based on the metrics of click speech, especially likes, imposes stress on

users. These new ramifications of communication additionally put stress on users on social media platforms without even being exposed to malicious actors in the first place. Hence, we do not only see further responsibilities for users but also further mental baggage regarding what users need to cope with. The dissertation also showed that coping strategies on social media do not account for all emotional states. Hence, for users, it is essential to take control of their communication habits and seek stress relief outside of the online information ecosystem. This can also happen by actively restricting online times.

8 Limitations

This dissertation has several limitations that go beyond the individual limitations of the papers. First, as outlined when it comes to information verification, users strongly rely on heuristics (**Papers II, III**, and **IV**). However, this dissertation does not provide extensive solutions on how social media platforms may implement mechanisms to increase or trigger critical thinking processes. There is discussion in the literature on several possibilities, such as ratings for sources (Kim et al. 2019) or inoculation, which acts as a sort of vaccine against misinformation (van Der Linden et al. 2020). Likewise, it does not propose a clear solution for the dilemma of balancing between the business model of social media platforms that seek maximum interaction and the reduction of extreme opinions.

The same issue applies to the spaces of social media communities. While the practical implications proposed ideas, the dissertation did not test any potential solutions to overcome the formation of echo chambers through supporting opinions against one's own beliefs (**Papers VI** and **VII**). As outlined, since users are now a crucial part of the information dissemination process, they have gained more responsibility in their information diet and their interaction habits. However, this dissertation does not show how such new responsibilities could effectively be communicated to the users in order to mitigate the unintended consequences. Existing literature on media literacy (Afrilyasanti et al. 2023) should focus strongly on the new role of users and the changes in communication patterns.

Concerning the solutions, there is also a big argument to be made whether governments should intervene in pushing for change in social media platforms. This dissertation only partially engaged with such arguments (Paper I), but did not emphasize how governments should intervene.

Furthermore, the dissertation strongly builds on interaction on the basis of inherent mechanisms of social media, such as likes and comments. It is clear that beyond these mechanisms, there are further ones, such as up-voting and downvoting on Reddit. It is also clear that interaction is not necessarily a "one and done" behavior in the sense that users only comment once on a post. There can be elaborate discussions on content that this dissertation does not cover.

A further limitation is that the dissertation did not explicitly focus on different demographics or groups. There is a possibility that different mechanisms, such as the fear of isolation (**Papers VI** and **VII**), may emerge stronger or weaker among elderly or young people, or across different political ideologies. Also, differences in nationality and culture could lead to different effects. The same applies to coping strategies

(Papers IX and X). The research focused on a US sample in the quantitative studies, and therefore, we may observe differences in non-Western countries, such as China.

Lastly, there is limited discussion about emerging technologies, such as artificial intelligence, and their role going forward in communication. Deepfakes can deceive users even further (Ahmed et al. 2023), and also, the reliance on generative AI may impose new challenges in how we communicate. Lastly, the increased authenticity of chatbots within social media platforms imposes a challenge for users to differentiate between actual human actors and bots.

9 Future Research

The limitations of this dissertation open up future research venues. This section provides different perspectives on how research within the changing communication environment can advance IS research in this specific context.

Investigating communication patterns across platforms. As Paper I suggests, social media platforms are the new intermediaries and gatekeepers of information with their algorithmic information curation. However, users are not only interacting with content on one platform but are part of various different platforms and communities. The online information ecosystem is the sum of all the platforms on which users interact. Hence, future research should focus more strongly on how information is transmitted and how users interact across different platforms. Similar to the social media communities mentioned in this dissertation, there may also be a perception associated with different platforms. Users may view the user basis of Facebook differently from the user basis of X (formerly Twitter). Such perceptions can further influence how users interact with content.

Feasibility of governmental regulation of social media platforms. For regulation, there are still question marks on how feasible it is for governments to intervene in the communication and information curation process. **Paper I** taps into a few cases of governmental intervention in the good (i.e., goal to reduce misinformation) and in a bad way (i.e., weaponization to silence opposition). There is still scarce research on how users would perceive such an intervention and how platforms may react to stronger regulation. Hence, to this day, it is not clear what governmental intervention should look like and to what extent governments even should regulate platforms.

Identifying demographics that significantly shape user interaction. As outlined by the limitations, there is also a need for more research in terms of different demographics or groups. **Papers VI** and **VII** investigate social media communities as an entity with members of similar interests. However, based on the interests, there may be differences in how content is consumed and interacted with. For example, a highly political social media community may incorporate stronger beliefs on certain topics. This, in return, can affect the fear of isolation in a stronger way as indicated by the topic importance moderation effect. Age and gender may also play a role, as well as the difference between more individualistic or collectivistic societies (Rhee et al. 1995).

Sentiment analysis within social media interaction. When we look at interaction through commenting, there is a lot to be learned about how users interact through language. As Papers **IX** and **X** outlined, interaction is tightly knit to emotions. Future research can investigate how emotions shape interaction after experiencing a negative user interaction and whether, in such a case, coping strategies actually lead to a change in emotions within users. Further emotions beyond anger and sadness should be explored in order to draw a full picture of how users react to which interaction types. Lastly, the emergence of emotions can be based on personality types. Therefore, it would be fruitful to investigate which personalities usually have an easier time experiencing negative user interactions and which group is most vulnerable to experiencing negative emotions.

Emerging technologies have benefits and risks in terms of interaction. As already outlined in the case of the presidential election of 2016 (Boichak et al. 2018), bot networks are on the rise on social media platforms. Also, Meta envisions a future in which everybody can create AI-powered accounts within their social media platforms (Criddle and Murphy 2024). These cases open up the discussion about user interaction in the sense that communication again takes another step in the direction of users not only communicating with other users, but with AI users. In this regard, it is elementary to understand how users perceive such accounts. Future research should focus on whether users are willing to engage with accounts that are flagged as AI accounts or how users identify non-human actors. Furthermore, the ramifications of such flooding of AI accounts need to be studied. Especially in the formation of public opinion, it is not clear how people may engage with other users if they do not agree with their opinion. A scenario that could play out is that opposing opinions are just labelled as "bot-driven," and users will refuse to engage with others as they think they do not actually engage with another human being.

10 Conclusion

Social media platforms change communication in different ways. One major change is the empowerment of users. Their role as passive information consumers switches to a role that incorporates interaction and, therefore, further information dissemination. To understand changes in modern communication, one has to understand the new role of users and their interaction behavior. Hence, this work focuses on the interaction aspect of social media in terms of "liking", "commenting", and "sharing". The results show that with this empowerment, users also gain a certain responsibility to shape the online information ecosystem. They vastly rely on heuristics to judge content, and oftentimes interact based on social influences. The judgment of the reliability or veracity of content is diminished in social contexts. Especially the notion of click-speech shows to be a behavior vastly driven by social aspects, such as the fear of isolation. This work provides a foundation for further investigations in user interaction, and fully understands unintended consequences stemming from interaction, as stress. Beyond stress, literature shows that such unintended consequences can also be the vast dissemination of false information or harmful narratives. Hence, to fully understand the issues of modern-day communication is to understand the ordinary social media user.

References

Acquisti, A. 2004. "Privacy in Electronic Commerce and the Economics of Immediate Gratification," *Proceedings of the 5th ACM conference on Electronic commerce*, pp. 21-29.

Acquisti, A., and Grossklags, J. 2005. "Privacy and Rationality in Individual Decision Making," *IEEE security & privacy* (3:1), pp. 26-33.

Afrilyasanti, R., Basthomi, Y., and Zen, E. L. 2023. "Factual Vs. Fake News: Teachers' Lens on Critical Media Literacy Education in Efl Classes," *Journal of Interactive Media in Education* (1), p. 16.

Ahmed, S., Wang, Y. F., and Bee, A. W. T. 2023. "Adjusting News Accuracy Perceptions after Deepfakes Exposure: Evidence from a Non-Western Context," *Telematics and Informatics* (84), p. 9.

Alashoor, T., and Baskerville, R. 2015. "The Privacy Paradox: The Role of Cognitive Absorption in the Social Networking Activity," in: *Proceedings of the 36th Conference on Information Systems (ICIS)*. Texas, USA: AIS eLibrary.

Amichai-Hamburger, Y., Kingsbury, M., and Schneider, B. H. 2013. "Friendship: An Old Concept with a New Meaning?," *Computers in Human Behavior* (29:1), pp. 33-39.

Awad, N. F., and Krishnan, M. S. 2006. "The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to Be Profiled Online for Personalization," *MIS quarterly* (30:1), pp. 13-28.

Bagozzi, R. P., and Yi, Y. 2012. "Specification, Evaluation, and Interpretation of Structural Equation Models," *Journal of the academy of marketing science* (40), pp. 8-34.

Bandura, A., and Walters, R. H. 1977. *Social Learning Theory*. Englewood cliffs Prentice Hall.

Bansal, G., Weinschenk, A., and Assoc Informat, S. 2020. *Something Real About Fake News: The Role of Polarization and Social Media Mindfulness*. Atlanta: Assoc Information Systems.

Bapna, S., Benner, M. J., and Qiu, L. 2019. "Nurturing Online Communities: An Empirical Investigation," *MIS quarterly* (43:2), pp. 425-452.

Baron, R. M., and Kenny, D. A. 1986. "The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of personality and social psychology* (51:6), p. 1173.

Barrera Jr, M. 1986. "Distinctions between Social Support Concepts, Measures, and Models," *American journal of community psychology* (14:4), pp. 413-445.

Bennett, W. L., and Segerberg, A. 2012. "The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics," *Information, communication & society* (15:5), pp. 739-768.

Bessi, A. 2016. "Personality Traits and Echo Chambers on Facebook," *Computers in Human Behavior* (65), pp. 319-324.

Beyens, I., Frison, E., and Eggermont, S. 2016. "“I Don’t Want to Miss a Thing”: Adolescents’ Fear of Missing out and Its Relationship to Adolescents’ Social Needs, Facebook Use, and Facebook Related Stress," *Computers in human behavior* (64), pp. 1-8.

Boichak, O., Jackson, S., Hemsley, J., and Tanupabrungsun, S. 2018. "Automated Diffusion? Bots and Their Influence During the 2016 Us Presidential Election," *International conference on information*: Springer, pp. 17-26.

Bolin, G. 2012. "The Labour of Media Use: The Two Active Audiences," *Information, Communication & Society* (15:6), pp. 796-814.

Bro, P., and Wallberg, F. 2014. "Digital Gatekeeping: News Media Versus Social Media," *Digital journalism* (2:3), pp. 446-454.

Caplan, G. 1974. *Support Systems and Community Mental Health: Lectures on Concept Development*, Behavioral Publication.

Cassel, J. 1976. "The Contribution of the Social Environment to Host Resistance: The Fourth Wade Hampton Frost Lecture," *American journal of epidemiology* (104:2), pp. 107-123.

Chen, Y., and Xu, Y. 2021. "Exploring the Effect of Social Support and Empathy on User Engagement in Online Mental Health Communities," *International Journal of Environmental Research and Public Health* (18:13), p. 6855.

Cicchetti, D. V., and Sparrow, S. A. 1981. "Developing Criteria for Establishing Interrater Reliability of Specific Items: Applications to Assessment of Adaptive Behavior," *American Journal of Mental Deficiency* (86:2), pp. 127-137.

Cinelli, M., Cresci, S., Galeazzi, A., Quattrociocchi, W., and Tesconi, M. 2020. "The Limited Reach of Fake News on Twitter During 2019 European Elections," *Plos One* (15:6), p. 13.

Citron, D. K. 2019. *The National Security Challenge of Artificial Intelligence, Manipulated Media, and 'Deep Fakes*. Boston University.

Cobb, S. 1976. "Social Support as a Moderator of Life Stress," *Psychosomatic medicine* (38:5), pp. 300-314.

Colliander, J. 2019. ""This Is Fake News": Investigating the Role of Conformity to Other Users' Views When Commenting on and Spreading Disinformation in Social Media," *Computers in Human Behavior* (97), pp. 202-215.

Coscia, M., and Rossi, L. 2020. "Distortions of Political Bias in Crowdsourced Misinformation Flagging," *Journal of the Royal Society Interface* (17:167), p. 20200020.

Criddle, C., and Murphy, H. 2024. "Meta Envisages Social Media Filled with Ai-Generated Users," in: *Financial Times*. https://www.ft.com/content/91183cbb-50f9-464a-9d2e-96063825bfcf.

Dennis, A. R., Fuller, R. M., and Valacich, J. S. 2008. "Media, Tasks, and Communication Processes: A Theory of Media Synchronicity," *MIS quarterly* (32:3), pp. 575-600.

Dennis, A. R., and Kinney, S. T. 1998. "Testing Media Richness Theory in the New Media: The Effects of Cues, Feedback, and Task Equivocality," *Information Systems Research* (9:3), pp. 256-274.

Devine, P. G., Hirt, E. R., and Gehrke, E. M. 1990. "Diagnostic and Confirmation Strategies in Trait Hypothesis Testing," *Journal of Personality and Social Psychology* (58:6), p. 952.

Dinev, T., Bellotto, M., Hart, P., Russo, V., Serra, I., and Colautti, C. 2006. "Privacy Calculus Model in E-Commerce–a Study of Italy and the United States," *European Journal of Information Systems* (15:4), pp. 389-402.

Ferm, L. E. C., and Thaichon, P. 2023. "I Think, Therefore I Ignore: A Study on Disinformation's Credibility Perceptions and Sharing Intentions over Social Media," in: *Journal of Strategic Marketing*. pp. 1-17.

Festinger, L. 1962. *A Theory of Cognitive Dissonance*. Stanford university press.

Folkman, S., and Lazarus, R. S. 1985. "If It Changes It Must Be a Process: Study of Emotion and Coping During Three Stages of a College Examination," *Journal of personality and social psychology* (48:1), p. 150.

Folkman, S., Lazarus, R. S., Dunkel-Schetter, C., DeLongis, A., and Gruen, R. J. 1986. "Dynamics of a Stressful Encounter: Cognitive Appraisal, Coping, and Encounter Outcomes," *Journal of personality and social psychology* (50:5), p. 992.

Fox, J., and Holt, L. F. 2018. "Fear of Isolation and Perceived Affordances: The Spiral of Silence on Social Networking Sites Regarding Police Discrimination," *Mass Communication and Society* (21:5), pp. 533-554.

Gabbiadini, A., Mari, S., and Volpato, C. 2013. "Virtual Users Support Forum: Do Community Members Really Want to Help You?," *Cyberpsychology, Behavior, and Social Networking* (16:4), pp. 285-292.

George, J. F., Gupta, M., Giordano, G., Mills, A. M., Tennant, V. M., and Lewis, C. C. 2018. "The Effects of Communication Media and Culture on Deception Detection Accuracy," *MIS quarterly* (42:2), pp. 551-575.

Gross, R., and Acquisti, A. 2005. "Information Revelation and Privacy in Online Social Networks," in: *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*. Alexandria, VA, USA: Association for Computing Machinery, pp. 71–80.

Guzman, A. L., and Lewis, S. C. 2020. "Artificial Intelligence and Communication: A Human–Machine Communication Research Agenda," *New media & society* (22:1), pp. 70-86.

Hair, J. F., Sarstedt, M., Ringle, C. M., and Mena, J. A. 2012. "An Assessment of the Use of Partial Least Squares Structural Equation Modeling in Marketing Research," *Journal of the academy of marketing science* (40), pp. 414-433.

Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., and Sarstedt, M. 2014. *A Primer on Partial Least Squares Structural Equation Modeling (Pls-Sem)*. SAGE Publications.

Halbach, P., Burbach, L., Nakayama, J., Plettenberg, N., Ziefle, M., Valdez, A. C., and Ieee. 2019. "Would I Lie to You? How Users Evaluate Faked Online Content Depending on Its Publication Type," in *2019 Ieee International Professional Communication Conference*. pp. 56-64.

Hayes, A. F., Matthes, J., and Eveland Jr, W. P. 2013. "Stimulating the Quasi-Statistical Organ: Fear of Social Isolation Motivates the Quest for Knowledge of the Opinion Climate," *Communication Research* (40:4), pp. 439-462.

Henseler, J., Ringle, C. M., and Sarstedt, M. 2015. "A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling," *Journal of the academy of marketing science* (43), pp. 115-135.

Hrynowski, Z. 2020. "Several Issues Tie as Most Important in 2020 Election," in: *Gallup*.

Jin, S. V., Muqaddam, A., and Ryu, E. 2019. "Instafamous and Social Media Influencer Marketing," *Marketing Intelligence & Planning* (37:5), pp. 567-579.

Johnsen, M. C. 2001. "Social Support Measurement and Intervention: A Guide for Health and Social Scientists," *Psychiatric services* (52:10), pp. 1404-1404.

Kahneman, D., Slovic, S. P., Slovic, P., and Tversky, A. 1982. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge university press.

Karahanna, E., Xu, S. X., Xu, Y., and Zhang, N. A. 2018. "The Needs–Affordances–Features Perspective for the Use of Social Media," *Mis Quarterly* (42:3), pp. 737-756.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., and Silvestre, B. S. 2011. "Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media," *Business horizons* (54:3), pp. 241-251.

Kim, A., and Dennis, A. R. 2019. "Says Who? The Effects of Presentation Format and Source Rating on Fake News in Social Media," *MIS Quarterly* (43:3), pp. 1025-1039.

Kim, A., Moravec, P. L., and Dennis, A. R. 2019. "Combating Fake News on Social Media with Source Ratings: The Effects of User and Expert Reputation Ratings," *Journal of Management Information Systems* (36:3), pp. 931-968.

Kitchens, B., Johnson, S. L., and Gray, P. 2020. "Understanding Echo Chambers and Filter Bubbles: The Impact of Social Media on Diversification and Partisan Shifts in News Consumption," *MIS quarterly* (44:4), pp. 1619-1649.

Klopfer, P. H., and Rubenstein, D. I. 1977. "The Concept Privacy and Its Biological Basis," *Journal of social Issues* (33:3), pp. 52-65.

Koriat, A., Lichtenstein, S., and Fischhoff, B. 1980. "Reasons for Confidence," *Journal of Experimental Psychology: Human learning and memory* (6:2), p. 107.

Krasnova, H., Spiekermann, S., Koroleva, K., and Hildebrand, T. 2010. "Online Social Networks: Why We Disclose," *Journal of information technology* (25:2), pp. 109-125.

Krumpal, I. 2013. "Determinants of Social Desirability Bias in Sensitive Surveys: A Literature Review," *Quality & quantity* (47:4), pp. 2025-2047.

Kushin, M. J., Yamamoto, M., and Dalisay, F. 2019. "Societal Majority, Facebook, and the Spiral of Silence in the 2016 Us Presidential Election," *Social Media+ Society* (5:2).

Landis, J. R., and Koch, G. G. 1977. "The Measurement of Observer Agreement for Categorical Data," *Biometrics*), pp. 159-174.

Larson, K., and Watson, R. 2011. "The Value of Social Media: Toward Measuring Social Media Strategies," *32nd International Conference on Information Systems (ICIS)*, Shanghai, China.

Lee, C. S., and Ma, L. 2012. "News Sharing in Social Media: The Effect of Gratifications and Prior Experience," *Computers in human behavior* (28:2), pp. 331-339.

Lewis, S. C., and Westlund, O. 2015. "Actors, Actants, Audiences, and Activities in Cross-Media News Work: A Matrix and a Research Agenda," *Digital journalism* (3:1), pp. 19-37.

Liang, H., Xue, Y., Pinsonneault, A., and Wu, Y. A. 2019. "What Users Do Besides Problem-Focused Coping When Facing It Security Threats: An Emotion-Focused Coping Perspective," *MIS Quarterly* (43:2), pp. 373-394.

Linvill, D. L., Boatwright, B. C., Grant, W. J., and Warren, P. L. 2019. ""The Russians Are Hacking My Brain!" Investigating Russia's Internet Research Agency Twitter Tactics During the 2016 United States Presidential Campaign," *Computers in Human Behavior* (99), pp. 292-300.

Linvill, D. L., and Warren, P. L. 2020. "Troll Factories: Manufacturing Specialized Disinformation on Twitter," *Political Communication* (37:4), pp. 447-467.

Lowry, P. B., D'Arcy, J., Hammer, B., and Moody, G. D. 2016. ""Cargo Cult" Science in Traditional Organization and Information Systems Survey Research: A Case for Using Nontraditional Methods of Data Collection, Including Mechanical Turk and Online Panels," *The Journal of Strategic Information Systems* (25:3), pp. 232-240.

Maier, C., Laumer, S., Eckhardt, A., and Weitzel, T. 2015. "Giving Too Much Social Support: Social Overload on Social Networking Sites," *European Journal of Information Systems* (24:5), pp. 447-464.

Malhotra, N. K., Kim, S. S., and Agarwal, J. 2004. "Internet Users' Information Privacy Concerns (Iuipc): The Construct, the Scale, and a Causal Model," *Information systems research* (15:4), pp. 336-355.

Maweu, J. M. 2019. ""Fake Elections"? Cyber Propaganda, Disinformation and the 2017 General Elections in Kenya," *African Journalism Studies* (40:4), pp. 62-76.

McCarthy, D., and Saegert, S. 1978. "Residential Density, Social Overload, and Social Withdrawal," *Human Ecology* (6:3), pp. 253-272.

McQuail, D., and Windahl, S. 2015. *Communication Models for the Study of Mass Communications*. Routledge.

Milan, S. 2015. "When Algorithms Shape Collective Action: Social Media and the Dynamics of Cloud Protesting," *Social Media+ Society* (1:2), p. 2056305115622481.

Moore, G. C., and Benbasat, I. 1991. "Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation," *Information Systems Research* (2:3), pp. 192-222.

Morales-i-Gras, J. 2020. "Cognitive Biases in Link Sharing Behavior and How to Get Rid of Them: Evidence from the 2019 Spanish General Election Twitter Conversation," *Social Media + Society* (6:2), p. 4.

Moravec, P., Minas, R., and Dennis, A. R. 2019. "Fake News on Social Media: People Believe What They Want to Believe When It Makes No Sense at All," *MIS Quarterly* (43:4), pp. 1343-1360.

Moy, P., Domke, D., and Stamm, K. 2001. "The Spiral of Silence and Public Opinion on Affirmative Action," *Journalism & Mass Communication Quarterly* (78:1), pp. 7-25.

Nabi, R. L., Prestin, A., and So, J. 2013. "Facebook Friends with (Health) Benefits? Exploring Social Network Site Use and Perceptions of Social Support, Stress, and Well-Being," *Cyberpsychology, behavior, and social networking* (16:10), pp. 721-727.

Neubaum, G., Rösner, L., Rosenthal-von der Pütten, A. M., and Krämer, N. C. 2014. "Psychosocial Functions of Social Media Usage in a Disaster Situation: A Multi-Methodological Approach," *Computers in Human Behavior* (34), pp. 28-38.

Neuwirth, K., Frederick, E., and Mayo, C. 2007. "The Spiral of Silence and Fear of Isolation," *Journal of Communication* (57:3), pp. 450-468.

Newcomb, T. M. 1953. "An Approach to the Study of Communicative Acts," *Psychological review* (60:6), p. 393.

Noelle-Neumann, E. 1977. "Turbulences in the Climate of Opinion: Methodological Applications of the Spiral of Silence Theory," *Public Opinion Quarterly* (41:2), pp. 143-158.

Noelle-Neumann, E. 1993. *The Spiral of Silence: Public Opinion - Our Social Skin*. University of Chicago Press.

Noelle-Neumann, E. 1974. "The Spiral of Silence a Theory of Public Opinion," *Journal of Communication* (24:2), pp. 43-51.

Norberg, P. A., Horne, D. R., and Horne, D. A. 2007. "The Privacy Paradox: Personal Information Disclosure Intentions Versus Behaviors," *Journal of consumer affairs* (41:1), pp. 100-126.

Pang, N., Ho, S. S., Zhang, A. M., Ko, J. S., Low, W., and Tan, K. S. 2016. "Can Spiral of Silence and Civility Predict Click Speech on Facebook?," *Computers in Human Behavior* (64), pp. 898-905.

Parameswaran, S., and Kishore, R. 2018. "Social Support in Online Health Communities: A Social-Network Approach," in: *SIGMIS-CPR'18: Proceedings of the 2018 ACM SIGMIS Conference on Computers and People Research*. pp. 93-94.

Park, I., and Lee, D. 2020. "Understanding News-Sharing across Different Platforms: The Effects of Newsworthiness and Gratifications from News-Sharing," *Information Research-an International Electronic Journal* (25:4).

Posner, R. A. 1981. "The Economics of Privacy," *The American economic review* (71:2), pp. 405-409.

Przybylski, A. K., Murayama, K., DeHaan, C. R., and Gladwell, V. 2013. "Motivational, Emotional, and Behavioral Correlates of Fear of Missing Out," *Computers in human behavior* (29:4), pp. 1841-1848.

Ragu-Nathan, T., Tarafdar, M., Ragu-Nathan, B. S., and Tu, Q. 2008. "The Consequences of Technostress for End Users in Organizations: Conceptual Development and Empirical Validation," *Information systems research* (19:4), pp. 417-433.

Rhee, E., Uleman, J. S., Lee, H. K., and Roman, R. J. 1995. "Spontaneous Self-Descriptions and Ethnic Identities in Individualistic and Collectivistic Cultures," *Journal of personality and social psychology* (69:1), pp. 142–152.

Robbins, I. P. 2014. "What Is the Meaning of Like: The First Amendment Implications of Social Media Expression," *Fed. Cts. L. Rev.* (7), p. 123.

Rogerson, P. A. 2019. *Statistical Methods for Geography: A Student's Guide*. Sage Publications.

Ross, B., Pilz, L., Cabrera, B., Brachten, F., Neubaum, G., and Stieglitz, S. 2019. "Are Social Bots a Real Threat? An Agent-Based Model of the Spiral of Silence to Analyse the Impact of Manipulative Actors in Social Networks," *European Journal of Information Systems* (28:4), pp. 394-412.

Ruijten, P. A. 2021. "The Similarity-Attraction Paradigm in Persuasive Technology: Effects of System and User Personality on Evaluations and Persuasiveness of an Interactive System," *Behaviour & Information Technology* (40:8), pp. 734-746.

Sahelices-Pinto, C., and Rodríguez-Santos, C. 2014. "E-Wom and 2.0 Opinion Leaders," *Journal of Food Products Marketing* (20:3), pp. 244-261.

Scheufele, D. A., Shanahan, J., and Lee, E. 2001. "Real Talk: Manipulating the Dependent Variable in Spiral of Silence Research," *Communication Research* (28:3), pp. 304-324.

Schuetz, S. W., Sykes, T. A., and Venkatesh, V. 2021. "Combating Covid-19 Fake News on Social Media through Fact Checking: Antecedents and Consequences," *European Journal of Information Systems* (30:4), pp. 376-388.

Seidman, G., Shrout, P. E., and Bolger, N. 2006. "Why Is Enacted Social Support Associated with Increased Distress? Using Simulation to Test Two Possible Sources of Spuriousness," *Personality and Social Psychology Bulletin* (32:1), pp. 52-65.

Selye, H. 1978. *The Stress of Life, Rev*. McGraw Hill.

Senaweera, M., Dissanayake, R., Chamindi, N., Shyamlal, A., Elvitigala, C., Horawalavithana, S., Wijesekera, P., Gunawardana, K., Wickramasinghe, M., Keppitiyagama, C., and Ieee. 2018. "The Influence of Community Interactions on User Affinity in Social Networks: A Facebook Case Study," in: *2018 National Information Technology Conference*.

Shannon, C. E. 2001. "A Mathematical Theory of Communication," *ACM SIGMOBILE mobile computing and communications review* (5:1), pp. 3-55.

Shore, J., Baek, J., and Dellarocas, C. 2018. "Network Structure and Patterns of Information Diversity on Twitter," *MIS quarterly* (42:3), pp. 849-A810.

Simon, H. A. 1979. "Rational Decision Making in Business Organizations," *The American economic review* (69:4), pp. 493-513.

Sklan, A. D. 2013. "@ Socialmedia: Speech with a Click of a Button-# Socialsharingbuttons," *Cardozo Arts & Ent. LJ* (32), p. 377.

Smith, H. J., Dinev, T., and Xu, H. 2011. "Information Privacy Research: An Interdisciplinary Review," *MIS quarterly* (35:4), pp. 989-1015.

Starbird, K., DiResta, R., and DeButts, M. 2023. "Influence and Improvisation: Participatory Disinformation During the 2020 Us Election," *Social Media+ Society* (9:2), p. 20563051231177943.

Strauss, A., and Corbin, J. 1990. *Basics of Qualitative Research*. Sage publications.

Stsiampkouskaya, K., Joinson, A., Piwek, L., and Ahlbom, C.-P. 2021. "Emotional Responses to Likes and Comments Regulate Posting Frequency and Content Change Behaviour on Social Media: An Experimental Study and Mediation Model," *Computers in Human Behavior* (124), p. 106940.

Sun, H. 2013. "A Longitudinal Study of Herd Behavior in the Adoption and Continued Use of Technology," *Mis Quarterly* (37:4), pp. 1013-1041.

Sun, R. Y., Li, C., Millet, B., Ali, K. I., and Petit, J. 2022. "Sharing News with Online Friends: A Study of Network Homophily, Network Size, and News Type," *Telematics and Informatics* (67).

Thatcher, J. B., Wright, R. T., Sun, H., Zagenczyk, T. J., and Klein, R. 2018. "Mindfulness in Information Technology Use: Definitions, Distinctions, and a New Measure," *MIS Quarterly* (42:3), pp. 831-848.

Thoits, P. A. 1995. "Stress, Coping, and Social Support Processes: Where Are We? What Next?," *Journal of health and social behavior*), pp. 53-79.

Top-Hashtags. 2023. "Top-Hashtags of Instagram." from https://top-hashtags.com/instagram/

Treem, J. W., and Leonardi, P. M. 2012. "Social Media Use in Organizations: Exploring the Affordances of Visibility, Editability, Persistence, and Association," *Communication yearbook* (36), pp. 143-189.

Tufekci, Z. 2014. "Engineering the Public: Big Data, Surveillance and Computational Politics," *First Monday* (19:7).

Vaast, E., Safadi, H., Lapointe, L., and Negoita, B. 2017. "Social Media Affordances for Connective Action: An Examination of Microblogging Use During the Gulf of Mexico Oil Spill," *MIS quarterly* (41:4), pp. 1179-1206.

Van Aelst, P., Van Erkel, P., D'heer, E., and Harder, R. A. 2017. "Who Is Leading the Campaign Charts? Comparing Individual Popularity on Old and New Media," *Information, communication & society* (20:5), pp. 715-732.

van Der Linden, S., Roozenbeek, J., and Compton, J. 2020. "Inoculating against Fake News About Covid-19," *Frontiers in Psychology* (11), p. 7.

Van Dijck, J. 2009. "Users Like You? Theorizing Agency in User-Generated Content," *Media, culture & society* (31:1), pp. 41-58.

Van Ingen, E., Utz, S., and Toepoel, V. 2016. "Online Coping after Negative Life Events: Measurement, Prevalence, and Relation with Internet Activities and Well-Being," *Social Science Computer Review* (34:5), pp. 511-529.

Venkatesan, S., Valecha, R., Yaraghi, N., Oh, O., and Rao, H. R. 2021. "Influence in Social Media: An Investigation of Tweets Spanning the 2011 Egyptian Revolution," *MIS Quarterly* (45:4), pp. 1679-1714.

Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. 2003. "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly* (27:3), pp. 425-478.

Watson, D., and Friend, R. 1969. "Measurement of Social-Evaluative Anxiety," *Journal of Consulting and Clinical Psychology* (33:4), p. 448.

Webster, J., and Watson, R. T. 2002. "Analyzing the Past to Prepare for the Future: Writing a Literature Review," *MIS quarterly* (26:2), pp. xiii-xxiii.

Wei, X., Zhang, Z., Zhang, M., Chen, W., and Zeng, D. D. 2019. "Combining Crowd and Machine Intelligence to Detect False News on Social Media," *MIS Quarterly* (46:2), pp. 977-1008.

Westley, B. H., and MacLean Jr, M. S. 1957. "A Conceptual Model for Communications Research," *Journalism Quarterly* (34:1), pp. 31-38.

White, D. M. 1950. "The "Gate Keeper": A Case Study in the Selection of News," *Journalism quarterly* (27:4), pp. 383-390.

Wirth, J., Maier, C., and Laumer, S. 2018. "The Influence of Resignation on the Privacy Calculus in the Context of Social Networking Sites: An Empirical Analysis," *Proceedings of the Twenty-Sixth European Conference on Information Systems*, Stockholm: Portsmouth, UK.

Wolfers, L. N., and Utz, S. 2022. "Social Media Use, Stress, and Coping," *Current Opinion in Psychology* (45), p. 101305.

Wolfswinkel, J. F., Furtmueller, E., and Wilderom, C. P. 2013. "Using Grounded Theory as a Method for Rigorously Reviewing Literature," *European journal of information systems* (22:1), pp. 45-55.

Wu, T.-Y., and Atkin, D. J. 2018. "To Comment or Not to Comment: Examining the Influences of Anonymity and Social Support on One's Willingness to Express in Online News Discussions," *New Media & Society* (20:12), pp. 4512-4532.

Wu, T.-Y., Oeldorf-Hirsch, A., and Atkin, D. 2020. "A Click Is Worth a Thousand Words: Probing the Predictors of Using Click Speech for Online Opinion Expression," *International Journal of Communication* (14), p. 20.

Xiao, X. Z. 2022. "Not Doomed: Examining the Path from Misinformation Exposure to Verification and Correction in the Context of Covid-19 Pandemic," *Telematics and Informatics* (74).

Yin, R. K. 2009. *Case Study Research: Design and Methods*. Sage Publications.

Zuckerberg, M. 2018. "A Blueprint for Content Governance and Enforcement." *Facebook Notes* Retrieved July 12, 2021, 2021, from https://www.facebook.com/notes/751449002072082/

Chapter I
Changes of Communication

Paper I

Information Flow, Gatekeeping, and Fake News:

Towards an Integrative Model for Social Media Communication

Maximilian Haug
Neu-Ulm University of Applied Sciences

Marten Risius
Neu-Ulm University of Applied Sciences

Heiko Gewald
Neu-Ulm University of Applied Sciences

Jason Bennett Thatcher
University of Colorado Boulder

Unpublished

Information Flow, Gatekeeping, and Fake News:

Towards an Integrative Model for Social Media Communication

Abstract

Social media has fundamentally altered communication roles and patterns within the online information ecosystem. In particular, social media platforms empower individual users to share information with large audiences, remove journalistic quality control barriers to producing news, and algorithmically curate content that is aligned with users' preexisting beliefs. However, these changes in information flow have led to various undesirable phenomena, including the broad dissemination of Fake News. While research has produced an extensive body of knowledge on the antecedents, drivers, and outcomes of such phenomena, we lack an integrated understanding of how the online information ecosystem promotes and mitigates such issues. In this article, we perform a grounded theory-based literature review of 665 Fake News articles published from 2014 to 2024 across disciplines to derive the Model of Online Networked Communication (MONC). The proposed model allows us to conceptualize the online information ecosystem's information flow. This article offers valuable insights into the communication roles and patterns found in the online information ecosystem's information flow and summarizes current knowledge on Fake News focused on the facilitating and mitigating factors.

Keywords

social media, Model of Online Networked Communication (MONC), information flow, grounded theory literature review method, Fake News, online information ecosystem, communication patterns, communication roles

Introduction

Social media platforms such as Facebook reportedly aim to "dominate" and "crush" traditional news industries by leveraging control over online information ecosystems—interconnected digital networks where information is created, disseminated, and consumed (Lima-Strong 2025). Countries such as Canada (Duggan 2023) and Australia (Kaye and Jackson 2024) recently enacted legislation requiring social media platforms to compensate news publishers for content. These regulatory responses underscore platforms' significant influence on public discourse, particularly through their role in disseminating Fake News—unverified or intentionally misleading content—highlighting a need to update theoretical frameworks that adequately capture these digitally driven information flows.

Social media significantly differs from traditional media by enabling simultaneous and interactive engagement among diverse participants (Dennis et al. 2008; Dennis and Kinney 1998). It involves complex interconnected networks of actors including platform operators, creators, media organizations, and users (Jin et al. 2019). Within these networks (e.g., news media providing articles for their audience), social media platforms act as technological gatekeepers through automated algorithms, governing information visibility and flow at unprecedented scales and speeds (Bro and Wallberg 2014).

Historically, gatekeeping—the filtering and control of information before dissemination—was performed by editors and fact-checkers at news organizations (Westley and MacLean Jr 1957; White 1950). Today, social media algorithms have largely supplanted these human gatekeepers, automatically curating and prioritizing content visibility (Kitchens et al. 2020; Lewis and Westlund 2015). This algorithmic governance significantly transforms user interactions with information, shifting communication from controlled editorial processes to automated, engagement-driven algorithms.

These fundamental gatekeeping changes require rethinking traditional audience roles. Platforms such as YouTube, Facebook, and Twitter (X) now function as primary intermediaries, facilitating interactive communication rather than simply delivering content (Tandoc Jr and Vos 2016). Audiences increasingly act as active participants who dynamically shape the flow of information through comments, shares, and likes. Traditional theories predominantly conceptualized communication as a linear flow from senders to passive receivers, yet contemporary communication demands models recognizing interactive, multidirectional feedback mechanisms (Till 2021).

As a result, online information ecosystems exhibit a dual, "Janus-faced" quality. Minimal gatekeeping allows rapid information dissemination and mobilization (Salge et al. 2022). As a result, they simultaneously facilitates misinformation (unintentional falsehoods) and disinformation (intentionally deceptive content) (Lazer et al. 2018). "Fake News," intentionally deceptive content mimicking legitimate news, spreads virally, often outpacing corrective information, exacerbating polarization, radicalization, and mistrust (Chan and Fu 2017; Lazer et al. 2018; Lu et al. 2022; Srivastava et al. 2020).

Recent events highlight Fake News's real-world impacts. During COVID-19, online misinformation undermined vaccination efforts, negatively influencing public behaviors (Allen et al. 2024). Similarly, misinformation significantly shaped voter perceptions during recent U.S. presidential elections and polarized public opinion around the Ukraine-Russia conflict. Although platform-level moderation (fact-checking, content removal, algorithmic deprioritization) and user-level initiatives (media literacy education) partially mitigate misinformation (Allcott and Gentzkow 2017; Livingstone and Helsper 2010), Fake News continues to spread rapidly, underscoring the urgency for contemporary theoretical frameworks explaining its dissemination.

Information Systems (IS) researchers are particularly positioned to advance theoretical insights into these dynamics. Traditional theories, such as Media Synchronicity Theory, emphasize synchronous communication but were originally developed for linear, one-to-many media environments (Dennis et al. 2008; Shannon 2001; Westley and MacLean Jr 1957). IS research has often uncritically applied linear

models to social media contexts (e.g., George et al. 2018), failing to fully account for algorithmic mediation, user-generated content, and collective online actions (Bennett and Segerberg 2012; Citron 2019; Larson and Watson 2011; Milan 2015; Vaast et al. 2017). We stress the shift from passive consumption of the audience towards active participation shapes opinions. Messaging becomes increasingly powerful through interaction while engaging with influencers who feel like friends in an environment that promotes creators that reinforce existing beliefs (DiResta 2024).

To close this theoretical gap, we developed the Model of Online Networked Communication (MONC) to explicitly capture interactive, multidirectional, and algorithmically mediated dynamics unique to social media ecosystems. Specifically, we ask:

> *How do characteristics of social media-driven online information ecosystems facilitate Fake News dissemination, and what mechanisms within these ecosystems mitigate its spread?*

We systematically synthesized existing Fake News research (665 articles, 2014–2024) through a grounded theory-based literature review (Wolfswinkel et al. 2013), generating novel theoretical insights. MONC uniquely identifies novel roles, communication patterns, research gaps, and opportunities for theoretical and practical exploration of Fake News.

MONC explicitly addresses theoretical limitations inherent in traditional models like Westley and MacLean's linear model, which inadequately captures multidirectional interactions enabled by digital technologies and social media. Traditional models fail to explain rapid misinformation proliferation during crises, such as COVID-19, where misinformation spreads via user-driven redistribution rather than through traditional gatekeepers. In contrast, MONC captures these digitally induced dynamics—algorithmic mediation, active audience redistribution, and multidirectional opinion leader interactions—previously unaccounted for by existing frameworks.

Thus, MONC's theoretical advancements addresses critical shortcomings of earlier models, offering a robust framework uniquely suited to analyzing complex modern online communication phenomena. By clearly defining roles such as algorithmic intermediaries, active audiences, and regulatory environments, MONC enhances theoretical precision, enabling more accurate empirical research and policy formulation to address the contemporary challenge of Fake News.

Conceptual Background

Applying a Communication Perspective to Understand Social Media Communication

Traditional communication theories, such as the ones of Shannon (2001) and Westley and MacLean Jr (1957) are linear sender-receiver models which lack sufficient explanatory power for contemporary online contexts characterized by interactive, many-to-many communication. Consequently, new theoretical models explicitly designed for digital ecosystems are necessary to capture accurately the complexities of modern information exchange.

We build on Westley and MacLean Jr (1957) traditional communication model, which aimed to clarify mass communication by identifying key roles—senders, receivers, and gatekeepers—our study leverages its foundational insights while directly addressing its limitations (e.g.,Bolin 2012; Van Dijck 2009). For example, in contrast to traditional models that conceptualize audiences as passive recipients, we model them as active and engaging. Nevertheless, Westley and MacLean's framework remains valuable for clearly articulating traditional mass communication processes, making it a suitable starting point for explicitly contrasting and highlighting the distinct interactive and dynamic characteristics of social media.

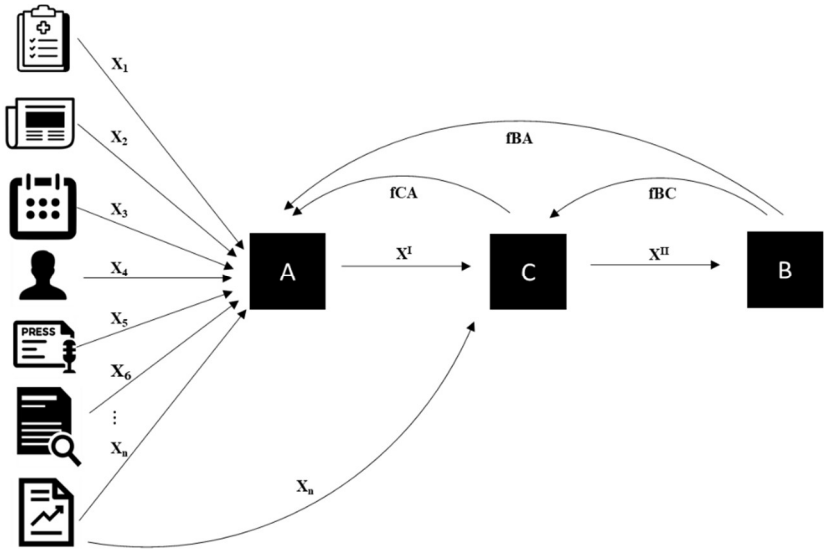


Figure 1. The Westley and MacLean Model of Communication

Westley and MacLean conceptualized communication as a process of selective filtering and shaping of information before dissemination to mass audiences. Actors A (e.g., political figures) or C (news media) choose specific information (sources: X1–Xn), disseminating it through mass media channels (e.g., television, radio, newspapers) to reach audiences (B) (McQuail and Windahl 2015). In this traditional configuration, news media act as gatekeepers, selectively transforming and mediating content for public consumption. Audience feedback occurs primarily through interpersonal channels, with limited scope for real-time engagement. While the model incorporates dynamic feedback loops (recipient–media, recipient–communicator, media–communicator), it predominantly emphasizes unidirectional communication flows, inadequately capturing today's interactive, networked media environments.

Contextualizing Westley and MacLean's model explicitly within contemporary social media environments reveals substantial transformations in communication dynamics. Although their basic sender–gatekeeper–receiver structure (A→C→B) maintains conceptual relevance, social media platforms fundamentally alter its mechanisms, primarily through algorithmically mediated content dissemination (Guzman and Lewis 2020; Lewis and Westlund 2015). For instance, platforms such as

Facebook and Twitter rely explicitly on automated recommendation algorithms to prioritize content, shaping the visibility and virality of information. This explicit algorithmic influence became especially apparent during the 2020 U.S. elections, where platform-driven content prioritization amplified polarizing narratives, significantly influencing voter engagement and public discourse.

Social media platforms thus explicitly disrupt traditional communication processes, creating largely unregulated interactive environments. Audience members have shifted explicitly from passive information consumers to active content creators who share, verify, and distribute information autonomously (Bansal et al. 2020; Sun et al. 2022; Xiao 2022). To explicitly capture and systematically explain these profound shifts in roles and interactions, we propose the Model of Online Networked Communication (MONC). MONC explicitly addresses theoretical gaps left by traditional models, clearly delineating evolving communication roles (e.g., active audiences, algorithmic intermediaries, influential opinion leaders) and novel interaction patterns (e.g., peer-driven message redistribution). By articulating these digitally induced transformations explicitly, MONC provides a comprehensive theoretical tool suited specifically for understanding modern social media dynamics.

Grounded Theory Literature Review Method

We employed the grounded theory literature review method (Wolfswinkel et al. 2013) to systematically derive the Model of Online Networked Communication (MONC) through an iterative analytical process. This structured, iterative process involved systematic open, axial, and selective coding, enabling the clear refinement and categorization of communication roles and patterns explicitly for contemporary social media contexts. Specifically, MONC emerged from six iterative analytical steps:

1. Initial model specification, explicitly drawing on traditional communication roles defined by Westley and MacLean Jr (1957).
2. Contextualization of roles, explicitly adapting traditional roles to social media contexts (e.g., transforming traditional gatekeepers into digital intermediaries).
3. Identification of many-to-many relationships, explicitly recognizing interactive dynamics characteristic of online environments (e.g., audience-to-audience interactions).
4. Reframing gatekeepers explicitly as intermediaries, clearly recognizing algorithmically mediated platforms (e.g., Facebook, Twitter) as primary channels of information dissemination.
5. Establishing model parsimony, explicitly reducing conceptual redundancy to clearly define distinct communication patterns.
6. Explicit integration of a regulatory environment, emphasizing critical governance mechanisms influencing online communication and information flow.

To rigorously ground MONC in contemporary empirical insights, we systematically leveraged an extensive body of peer-reviewed Fake News literature (665 articles, 2014–2024) identified through predefined keyword searches in leading journal databases (e.g., Web of Science). Fake News is explicitly defined as intentionally produced disinformation designed to mimic credible news (Lazer et al. 2018).

Empirical examples, such as misinformation during the COVID-19 pandemic and recent elections, explicitly illustrate how social media has profoundly disrupted traditional communication roles (e.g., opinion leaders, intermediaries) and audience dynamics, thereby underscoring the necessity of MONC.

Expanding explicitly upon Westley and MacLean Jr (1957) traditional model, MONC systematically identifies and categorizes technologically induced transformations in communication roles (e.g., traditional gatekeepers transitioning into opinion leaders) and interaction patterns (e.g., user-driven message redistribution). Additionally, MONC explicitly incorporates new roles critical in contemporary information ecosystems, such as digital intermediaries (social media platforms) and local regulatory environments (content moderation frameworks). This explicitly structured analytical process ensures MONC remains robustly grounded in contemporary misinformation research.

We explicitly operationalized communication roles within MONC as follows:

Events: occurrences or developments prompting communication.

Opinion Leaders: actors explicitly leveraging social support and influence to shape information dissemination (Katz and Lazarsfeld 1964).

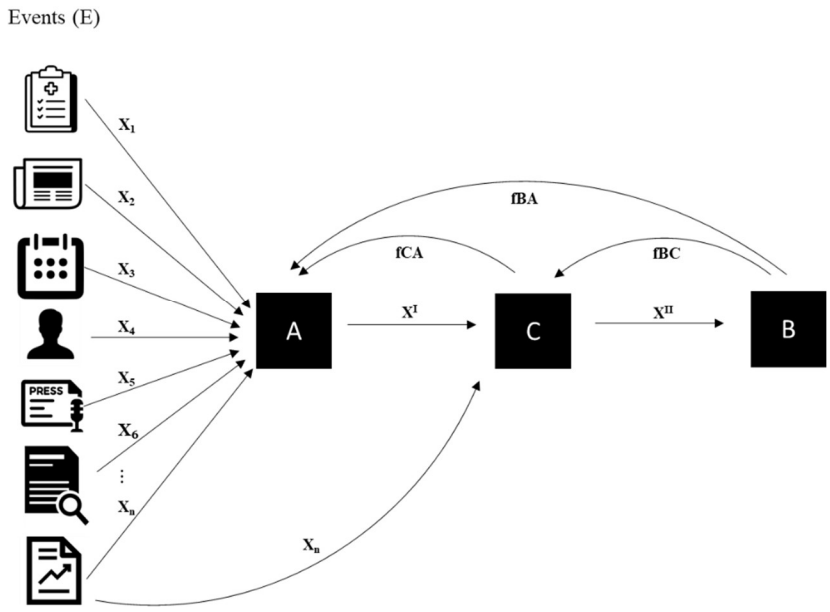
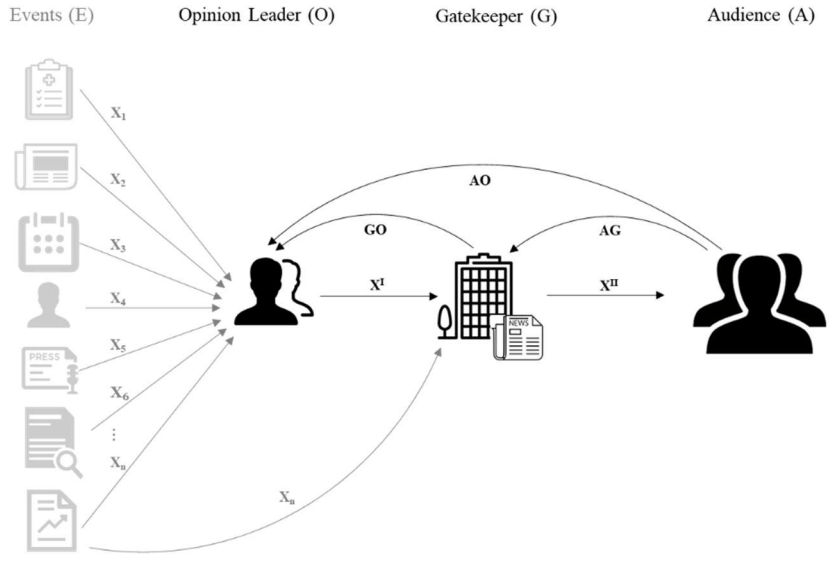
Intermediaries: digital platforms explicitly enabling user-generated content creation and exchange, thus serving as foundational communication infrastructures (Kaplan and Haenlein 2010).

Audience: individuals or groups explicitly engaging with information to fulfill specific communication needs (Westley and MacLean Jr 1957).

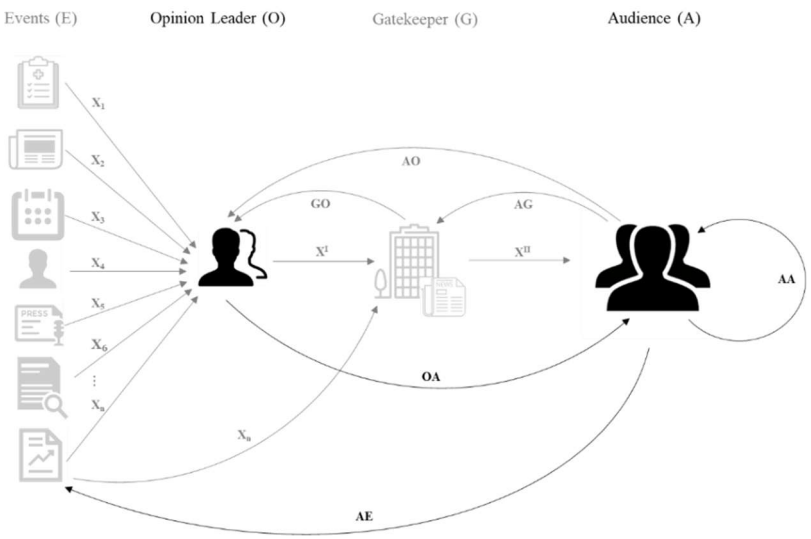
Regulatory Environment: legal frameworks and governance mechanisms explicitly regulating online communication and speech (Andorfer 2018).

Furthermore, our iterative coding explicitly recognized the dynamic fluidity of platform roles. Initially defined as intermediaries aligning with standard legislative definitions (European Parliament and of the Council 2022), platforms such as Facebook, Twitter, Instagram, or TikTok occasionally shift explicitly into gatekeeping roles, significantly exerting editorial influence over information flow (as defined by Westley and MacLean Jr (1957). Therefore, gatekeepers explicitly constitute a subset of intermediaries characterized by heightened control and influence over content visibility.

Finally, we operationalized communication patterns explicitly aligned with interactive, feedback-oriented processes described by Westley and MacLean Jr (1957), adapted explicitly to contemporary social media contexts through iterative grounded theory analysis. The subsequent section explicitly details each iterative step, clearly presenting MONC's evolution from initial theoretical formulation to its final comprehensive form.

Visualization	Implementations
Iteration #1: Specify Initial Model	
	Apply the communication model of Westley and MacLean Jr (1957) to explore the viability of communication roles and patterns in the Fake News literature.
Iteration #2: Contextualization of Communication Roles	
	In line with the initial role, we contextualized A, B, and C to the online communication context. **A -> Opinion Leader** (i.e., influencer, political figure) **C -> Gatekeeper** (i.e., new and traditional news organizations) **B -> Audience** (i.e., target recipients)

Iteration #3: Determine Many-to-Many Relationships



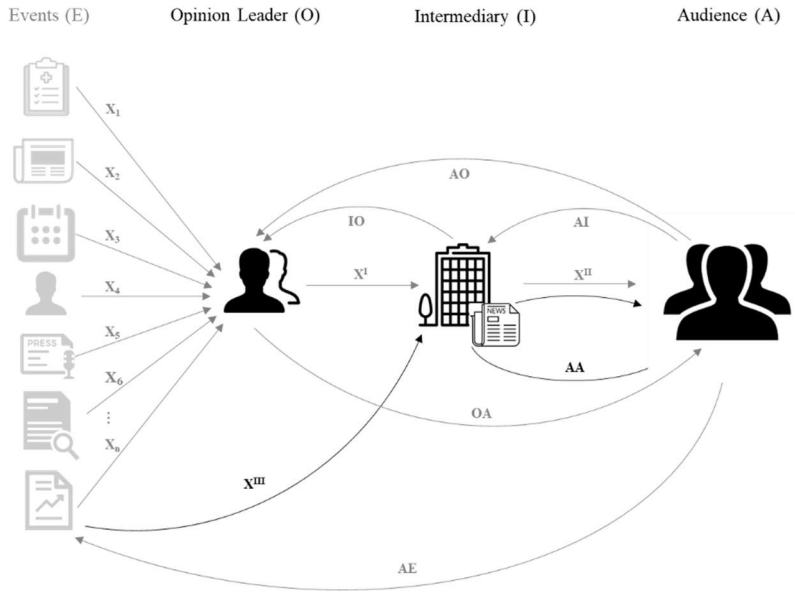
Within the online information ecosystem, three communication patterns emerge that create many-to-many relationships.

OA: Opinion leaders directly communicate with the audience

AA: Audience interacts with each other

AE: Audience is empowered to engage with source information

Iteration #4: Reframing of Gatekeepers to Intermediaries



Social media platforms emerge as the intermediary:

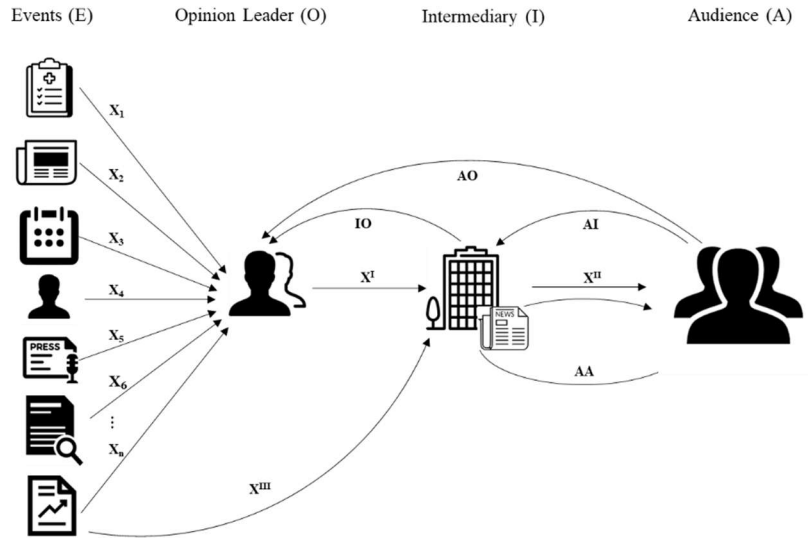
Intermediary: platforms become intermediary

Opinion Leader: news organizations become alternate opinion leaders

AA: communication pattern is now dependent on the platform intermediary.

X^{III}: platforms use ban lists to filter content proactively

Iteration #5: Establish Model Parsimony

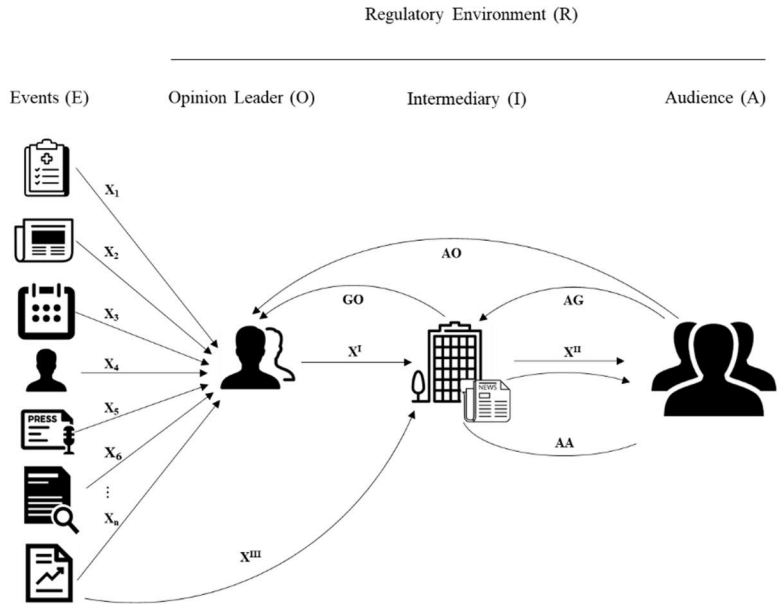


Parsimonious reduction of relationships

AE: engagement with events as part of audience characteristics, since the engagement alone is not a communication pattern

OA: communication of opinion leaders with the audience is disseminated through the intermediary (X^{I} ->X^{II})

Iteration #6: Add Regulatory Environment – Model of Online Networked Communication



Empirical-to-conceptual: Addition of **regulatory environment** in order to provide a full picture of the online information ecosystem. The regulatory environment affects opinion leaders, intermediaries, and the audience for what they are legally able to communicate and under which circumstances platforms are regulated.

Table 1. Evolution of Iterative Model Development

Notes: Descriptions of arrows after iteration #1 describe which role communicates with another role, e.g., AG means: audience communicates with an intermediary.

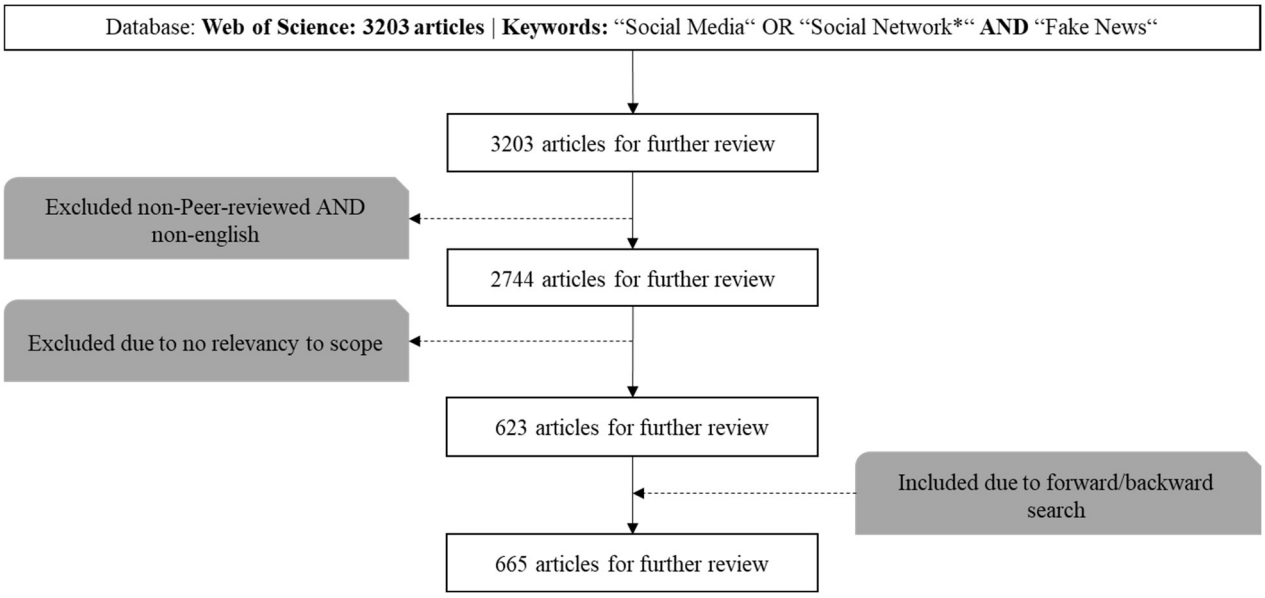


Figure 2. Identified Relevant Articles (665 Studies) Within the Scope of the Literature Review

Notes: Date of search 13th of April 2024, Database searched: Web of Science

Results

Model of Online Networked Communication (MONC)

Our six iterations of analysis resulted in the final version of MONC (Figure 3). Each iteration systematically refined MONC through coding and categorization, progressively clarifying and defining the roles and communication patterns within the model. This version captures the online information ecosystem through a combination of five roles (Events, Opinion Leaders, Intermediaries, Audience, and Regulatory Environment) and eight communication patterns (X1…n, X^{I}, X^{II}, IO, AI, AO, AA, and X^{III}) (Table 2 and Table 3, respectively). Illustrative examples derived from literature, such as misinformation during the COVID-19 pandemic and political polarization during recent elections, concretely informed these roles and communication patterns.

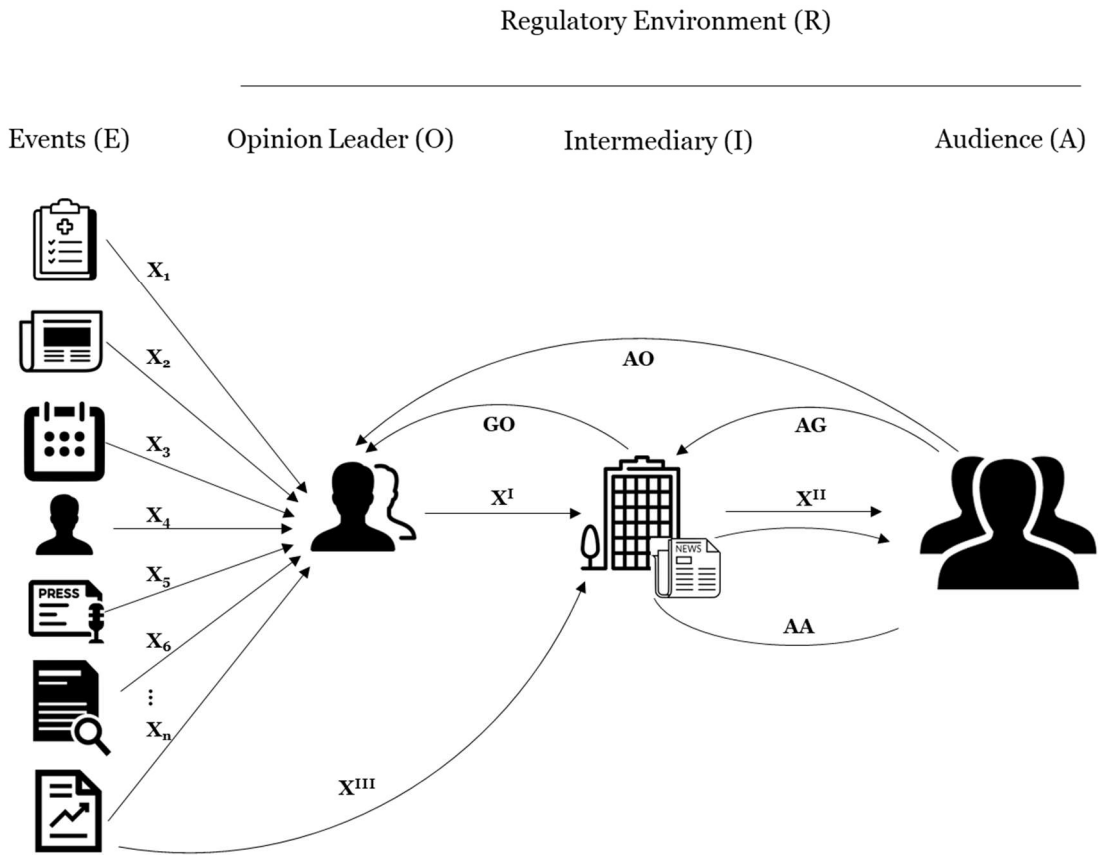


Figure 3. Full Version of the Model of Online Networked Communication (MONC)

Construct	Definition	Operationalization	Coding Examples
Communication roles			
Events (E)	The totality of objects and events "out there" (Westley and MacLean Jr 1957)	Subjects of communication that occur to inspire communication in the first place.	COVID-19 pandemic, US presidential elections, and event characteristics
Opinion leader (O)	Opinion leaders use their social support and social pressure to influence their respective networks (Durani et al. 2023; Katz and Lazarsfeld 1964).	Individuals or organizations act as opinion leaders whose accounts of events have a substantial influence on their audiences. The main characteristics of opinion leaders on social media include high following and/or high engagement with their content to drive public opinion.	Characteristics and types of opinion leaders, such as government officials, traditional news media, Fake News media, hyperpartisan media, and political influencers that push or refute Fake News
Intermediary (I)	Internet-based applications that allow the creation and exchange of user-generated content (Kaplan and Haenlein 2010)	Online platforms that allow users to create, share, and engage with curated user-generated content through endorsement mechanisms (e.g., likes, up-votes, and favorites) or text-based communication (e.g., comment, share, and retweet), build networks with others, and maintain a personal profile.	Features of social media platforms, such as sharing, commenting, and content recommender algorithms; platform policies, content moderation, and fact checking practices and organizations; and trust and safety teams that influence the impact of Fake News.
Audience (A)	A person or social system satisfying informational needs through communication (Westley and MacLean Jr 1957)	Individuals or groups who predominantly use online networks to acquire information and engage with content.	Characteristics of the audience, such as age, tech-savviness, or media literacy, that make them either susceptible or resistant to Fake News.
Regulatory environment (R)	Legal regulations governing speech and communication within the online information ecosystem (Andorfer 2018)	Legislature that affects the online information ecosystem (e.g., regional laws that regulate freedom of expression, define protected attributes, and attribute responsibilities for content moderation).	Policies and policy proposals (e.g., the First Amendment, Section 230 (USA), GDPR (Europe), and NetzDG (Germany)) that influence the spread and moderation of Fake News, along with censorship practices to mitigate Fake News.

Table 2. Communication Roles with Definitions and Examples

Construct	Definition	Operationalization	Coding Examples
Communication patterns			
X1…n	Messages containing information about an event.	Information in various formats (eyewitness reports, video footage, data) selected by opinion leaders to create narratives and share with their audiences.	Videos or posts documenting events like protests, natural disasters, or political developments circulated online.
X^{I}	Opinion leaders publicize online.	Opinion leaders select, contextualize, create, and share messages using communication platforms to target their audiences.	Political influencers crafting persuasive narratives mimicking credible news to spread Fake News or discredit opponents.
X^{II}	Intermediary curates social media content.	Platform content moderation and curation mechanisms influencing visibility and distribution of Fake News.	Algorithmic curation that reinforces users' exposure to previously viewed or similar Fake News content.
IO	Intermediary moderates opinion leaders.	Platform-based moderation affecting opinion leaders' visibility and reach.	Shadow-banning or deplatforming controversial opinion leaders due to spreading Fake News.
AI	Audience directly addresses the intermediary.	Audience actively fact-checks or reports misleading content directly to platforms.	Users reporting suspicious posts or Fake News due to distrust in specific opinion leaders or awareness of misinformation campaigns.
AO	Audience engages opinion leader via intermediary.	Audience reactions influenced by content disseminated by opinion leaders.	Audience interactions (comments, likes, criticisms) reflecting belief or skepticism towards Fake News shared by influential figures.
AA	Audience redistributes messages among themselves.	Active audience participation in the dissemination and discussion of misinformation within their social networks.	Instances where audiences amplify Fake News through sharing, driven by novelty, believability, emotional impact, or trust factors.
X^{III}	Intermediary proactively filters original content.	Platform-driven proactive filtering methods to prevent certain types of misinformation from appearing.	Using shared hash databases across platforms to block terrorist content or extreme Fake News before publication.

Table 3. Communication Patterns with Definitions and Examples

Key Digitally Induced Changes to the Online Information Ecosystem

Building on the operationalization of communication roles and patterns outlined previously, we turn to highlighting key technologically induced changes identified by MONC. Specifically, our iterative grounded theory analysis reveals four fundamental shifts that distinguish online information ecosystems from traditional communication environments. Specifically, MONC identifies: (1) social media platforms emerging as intermediaries, (2) news organizations transitioning into opinion leaders, (3) enhanced audience empowerment through engagement, and (4) the rising importance of regulatory frameworks. These shifts illustrate how social media platforms have redefined communication roles and dynamics, significantly reshaping how Fake News is facilitated or mitigated.

First, social media platforms act as intermediaries by providing infrastructure and communication channels for actors to connect with and engage target audiences (Fourney et al. 2017). They algorithmically curate and promote content (Balmas 2014) as well as moderate and restrict information flows (Moravec et al. 2018). Deviating from the traditional model of communication, MONC designates platforms as "intermediaries" rather than traditional "gatekeepers," reflecting nuanced distinctions recently introduced by regulatory frameworks to classify online platforms (European Commission 2023).

Second, platforms now serve dual roles, acting not only as intermediaries but also as influential opinion leaders, supplanting traditional gatekeepers such as news media organizations (Grossman 2022). MONC expands the traditional concept of advocacy (Westley and MacLean Jr 1957) to encompass actors who significantly influence audience perceptions and behaviors (Katz and Lazarsfeld 1964). As such, direct platform access to large user bases and interactive audience engagement have transformed opinion leadership (Parikh et al. 2019), fostering new types of influencers and online opinion leaders (Durmaz and Hengirmen 2022; Guo and Zhang 2020).

Third, social media platforms empower users by facilitating interactive many-to-many communication. Accordingly, MONC explicitly addresses how audiences actively redistribute information within their networks (AA) (Cinelli et al. 2020a; Sun et al. 2022) and directly engage with opinion leaders (AO) (Coscia and Rossi 2020), transforming users into proactive participants within information ecosystems.

Fourth, MONC emphasizes the increasingly pivotal role of regulatory frameworks established by governing bodies in shaping online information ecosystems. As platforms enable communication at global scales, they must navigate complex and diverse legal environments, particularly regarding significant societal issues amplified by social media, such as Fake News (Pang et al. 2022).

In applying MONC to synthesize the Fake News literature, we structured existing research into clear themes. Current Fake News research primarily emphasizes audience roles (A – 34.0%) and event characteristics (E – 18.2%), along with audience-driven redistribution of information (AA – 17.3%) and opinion leaders' dissemination strategies (XI – 15.2%). However, significant research gaps remain

regarding opinion leader characteristics (O – 7.8%), intermediaries' roles (I – 3.0%), regulatory environments (R – 4.0%), platform information filtering mechanisms (XII – 8.6%, XIII – 0.3%), and the effects of content moderation activities (AI – 2.4%, AO – 7.0%, IO – 1.3%).

Overall, current research underscores how specific communication roles and patterns either facilitate or mitigate Fake News. Here, 'facilitate' refers to processes involved in the creation, dissemination, and amplification of Fake News, while 'mitigate' refers to characteristics and mechanisms that inhibit or reduce its spread and impact. Figure 4 summarizes these findings, highlighting descriptive statistics and thematic insights across communication roles (events, opinion leaders, intermediaries, audience, regulatory environment) and patterns. The following sections provide further elaboration on each role and communication pattern, supported by representative sources and concrete examples.

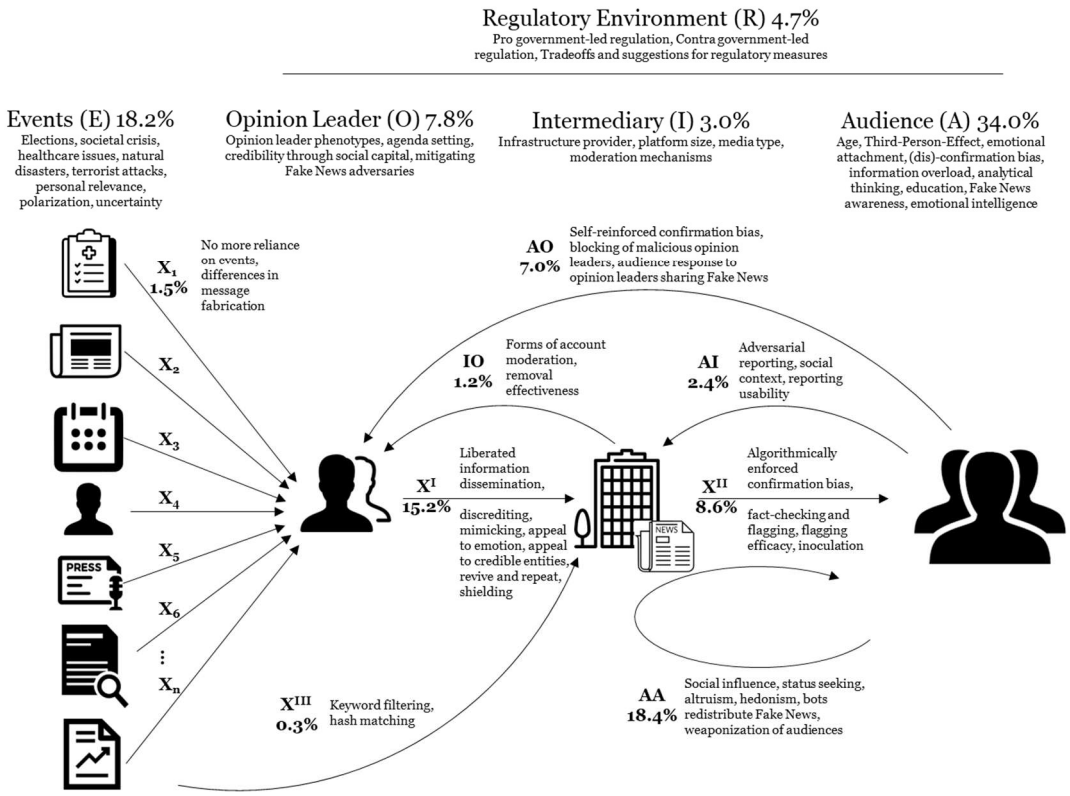


Figure 4. Overview of the Literature Review Results on Article Topics (N = 665)

Communication Roles and Fake News

Events

Events represent occurrences and topics that inspire communication (Westley and MacLean Jr 1957). MONC categorizes events by news subject (elections, societal crises, healthcare issues, natural disasters, terrorist attacks) and thematic characteristics (personal relevance, polarization, uncertainty) (see Table 4). Fake News research primarily focuses on political events such as elections, leveraging heightened public interest and polarization (Guess et al. 2020). Healthcare issues, notably COVID-19, are targeted

due to associated uncertainty (Gallotti et al. 2020). Societal crises, including movements like Black Lives Matter and conflicts like the war in Ukraine, are also significant contexts for Fake News (Morejon-Llamas et al. 2022). Conversely, natural disasters and terrorist attacks receive comparatively less attention (Kwanda and Lin 2020; Strand and Svensson 2019; Xu et al. 2020; Yan et al. 2021).

Theme	Findings	Exemplar Sources
Facilitation – News Subject		
Elections	Fake News predominantly targets electoral events, especially high-profile governmental elections (e.g., U.S. presidential elections), exploiting heightened public interest and polarization.	(Cinelli et al. 2020b; Guess et al. 2020; Lepird et al. 2024; Maweu 2019; Zimmermann and Kohring 2020)
Societal crisis	Fake News frequently emerges in response to ongoing societal crises that attract significant public attention. These crises range from localized political movements (e.g., Catalan independence, Brexit) to international conflicts (e.g., the war in Ukraine) and global social justice movements (e.g., Black Lives Matter, LGBTQI rights, women's rights protests in Iran).	(del Vas and Navarro 2024; Durani et al. 2023; Morejon-Llamas et al. 2022; Wenzel et al. 2023; Wilson and Starbird 2021)
Healthcare issues	Fake News predominantly addressed the COVID-19 pandemic, particularly vaccine misinformation. Other health issues, such as the Zika virus, received relatively less attention.	(Ahmed et al. 2020; Baumann and Humprecht 2024; Gallotti et al. 2020; Slutskiy and Boonchutima 2022; Wu et al. 2023)
Natural disasters	Fake News occasionally targets natural disasters, particularly earthquakes and tsunamis.	(Kwanda and Lin 2020; Méndez-Muros et al. 2024)
Terrorist attacks	Fake News arises during significant terrorist incidents, including events such as the Boston Marathon bombing and the 2017 UK terror attacks.	(Innes et al. 2021; Kostakos et al. 2018; Pantumsinchai 2018; Rosu et al. 2023)
Facilitation – Thematic Enablers		
Personal relevance	Fake News frequently exploits events likely to evoke personal relevance and emotional engagement, such as fabricated stories of individuals or animals in distress.	(Dumitru 2020; Grinberg et al. 2019; Levens et al. 2018; San Martin et al. 2020)
Polarization	Events characterized by high polarization potential provide fertile ground for Fake News, stimulating increased audience engagement and interest.	(Allcott and Gentzkow 2017; Prochaska et al. 2023; Wilson and Starbird 2021; Zimmermann et al. 2022)
Uncertainty	Events marked by uncertainty, such as novel pandemics, are prime targets for Fake News, as misinformation fills initial information gaps and meets audience information needs.	(Gallotti et al. 2020; Himelein-Wachowiak et al. 2021; Sommariva et al. 2018)

Table 4. How Events, Personal Relevance, Polarization, and Uncertainty Enable Fake News

Across these events, three thematic enablers facilitate Fake News: personal relevance, polarization, and uncertainty. Personal relevance refers to events directly impacting individuals, increasing susceptibility to misinformation (Levens et al. 2018). Polarization involves events that amplify societal divisions, especially elections (Prochaska et al. 2023). Uncertainty characterizes events with incomplete or evolving information, exemplified by the early stages of COVID-19, where misinformation fills information gaps (Gallotti et al. 2020).

Opinion Leaders

In the online information ecosystem, opinion leaders are influential individuals or organizations (e.g., news media, influencers, politicians, and firms) that actively shape public perceptions of events (Bergström and Jervelycke Belfrage 2018; Durani et al. 2023; Katz and Lazarsfeld 1964; Westley and MacLean Jr 1957). Opinion leaders who facilitate Fake News—termed "malicious opinion leaders"—include both real or fabricated media figures, politicians, government officials) (Durmaz and Hengirmen 2022; Guo and Zhang 2020), and real or Fake News media (e.g., Breitbart News Network) (Parikh et al. 2019). They effectively engage in agenda-setting by rapidly disseminating misinformation without rigorous journalistic oversight (Starbird et al. 2023), leveraging substantial credibility and social capital to amplify their influence and resist fact-checking (Sterrett et al. 2019; Zimmer and Reich 2018).

Theme	Findings	Exemplary Sources
Facilitation – Malicious Opinion Leaders' Forms of Operation		
Opinion leader phenotypes	Malicious opinion leaders consist of real or fabricated individuals (e.g., media personalities, politicians, officials) and media outlets disseminating Fake News.	(Durmaz and Hengirmen 2022; Guo and Zhang 2020; Parikh et al. 2019)
Agenda setting	Malicious opinion leaders actively set agendas by quickly disseminating Fake News due to a lack of journalistic accountability.	(Berry et al. 2023; Hamdi 2023; Lazer et al. 2018; Starbird et al. 2023)
Credibility through social capital	Malicious opinion leaders exploit their substantial social capital and credibility (e.g., high follower counts) to effectively spread Fake News and resist fact-checking efforts.	(Christner 2023; Flintham et al. 2018; Le 2024; Millet et al. 2024; Osatuyi and Dennis 2024; Sterrett et al. 2019; Zimmer and Reich 2018)
Mitigation – Journalistic Opinion Leaders as Opponents of Fake News		
Adversaries Mitigating Fake News	Journalistic opinion leaders, primarily established news organizations, actively counter Fake News by engaging in rigorous fact-checking and distinguishing their credibility from malicious sources.	(Esteban-Navarro et al. 2021; Krasni 2020; Vu and Saldaña 2021; Zecchinon and Standaert 2025)

Table 5. Opinion Leaders: Facilitation and Mitigation of Fake News

Conversely, news media organizations have transitioned from traditional gatekeeping to acting as journalistic opinion leaders, directly countering Fake News by explicitly positioning themselves as ideological adversaries of malicious opinion leaders. They achieve this by actively engaging in and advocating rigorous fact-checking initiatives, thereby reinforcing their adherence to high journalistic

standards and credibility (Esteban-Navarro et al. 2021), and adhering to rigorous journalistic standards (Krasni 2020).

Audience

The audience is comprised of recipients whose communication needs are satisfied by opinion leaders (Westley and MacLean Jr 1957). However, changes involving opinion leaders and intermediaries have increased audience vulnerability, making it more challenging to differentiate accurate information from Fake News (see e.g., Grossman 2022).

Our review identifies several audience characteristics facilitating susceptibility to Fake News: advanced age (Shu et al. 2018; Tantau et al. 2018), third-person-effect (i.e., the perception that only others fall for Fake News), third-person effect—perceiving only others as vulnerable (Cheng and Luo 2021), and heightened emotional attachment to the news subject (Xu et al. 2020) are characteristics that heighten the likelihood of believing Fake News. Additionally, cognitive factors such as confirmation and disconfirmation biases toward both content and sources (Pennycook et al. 2021; Tandoc et al. 2021; Turel 2023) and information overload (Tandoc and Kim 2022; Zhang et al. 2022a) further increase susceptibility to misinformation.

Conversely, several protective factors predict audience engagement in information verification behaviors, thereby mitigating susceptibility to Fake News. These include enhanced analytical thinking skills (Pennycook and Rand 2019a; Pennycook and Rand 2019b), higher education levels (Pop and Ene 2019; Rampersad and Althiyabi 2020), general awareness of Fake News (Moravec et al. 2022), and greater emotional intelligence (Preston et al. 2021; Wu et al. 2023).

Theme	Findings	Exemplary Sources
Facilitation – Audience Members' Vulnerabilities		
Age	Older individuals exhibit higher vulnerability to Fake News, although underlying mechanisms remain insufficiently explored.	(Rampersad and Althiyabi 2020; Shu et al. 2018; Tantau et al. 2018)
Third-person effect	Individuals commonly believe others are more susceptible to Fake News, reflecting overconfidence in their own discernment abilities.	(Cheng and Luo 2021; Chung and Kim 2021; Suttle et al. 2021; Tian and Willnat 2025)
Emotional attachment	Heightened emotional investment (e.g., personal relevance, polarization) increases susceptibility to Fake News.	(Dumitru 2020; Tan and Hsu 2023; Xu et al. 2020)
(Dis)-Confirmation bias	Audiences readily accept information aligning with their existing beliefs (confirmation bias) and dismiss contradictory information as Fake News (disconfirmation bias), influencing trust and verification behaviors.	(Pennycook et al. 2021; Tandoc et al. 2021; Turel 2023)
Information overload	Information overload increases vulnerability to Fake News; repeated exposure further reinforces belief in misinformation.	(Boulianne et al. 2022; Tandoc and Kim 2022; Zhang et al. 2022a)

Mitigation – Information Verification Predictor		
Analytical thinking	Analytical thinking enhances an individual's ability to distinguish credible from non-credible sources and increases responsiveness to fact-checking efforts.	(Acomi et al. 2021; Hu and Apuke 2023; Pennycook and Rand 2019a; Pennycook and Rand 2019b)
Education	Higher educational levels are linked to better analytical skills and stronger information verification behaviors.	(Buturoiu et al. 2021; Jiang and Wang 2024; Maningo et al. 2022; Miller et al. 2024; Pop and Ene 2019; Rampersad and Althiyabi 2020)
Fake News awareness	General awareness about Fake News enhances skepticism, reducing the likelihood of believing and sharing misinformation.	(Anspach and Carlson 2022; Daunt et al. 2023; Moravec et al. 2022; Omar et al. 2023; Verma and Nayak 2024)
Emotional intelligence	Emotional intelligence, through effective emotion management, reduces susceptibility to sensationalist Fake News content.	(Preston et al. 2021; Wu et al. 2023)

Table 6. Audience Vulnerabilities and Protective Factors Influencing Fake News Believability

Intermediaries

Social media platforms act as intermediaries by providing infrastructure enabling many-to-many interactions between opinion leaders and their audiences (Fourney et al. 2017). Two primary platform characteristics shape Fake News dissemination. First, platforms create an open ecosystem where content-sharing mechanisms—such as commenting, sharing, and recommendation—enable rapid dissemination of Fake News (Guo and Zhang 2020). Two primary platform characteristics shape Fake News dissemination. First, platforms create an open ecosystem where content-sharing mechanisms—such as commenting, sharing, and recommendation—enable rapid dissemination of Fake News (Balmas 2014) (see section XII for detailed insights on curation algorithms and filters). Such intermediary characteristics further affect Fake News facilitation or mitigation. (Fourney et al. 2017). Smaller platforms, characterized by heterogeneous content and limited moderation (e.g., Gab), particularly cater to hyperpartisan or fringe audiences, increasing polarization and facilitating Fake News dissemination (Peucker and Fisher 2022). Platform media types—textual (e.g., Twitter) versus visual (e.g., Instagram)—have limited direct impact on Fake News spread. (Valenzuela et al. 2022). However, platforms emphasizing closed, interpersonal communication (e.g., WhatsApp, Telegram) notably facilitate misinformation (Neyazi et al. 2022). To mitigate Fake News, platforms employ varying moderation strategies, such as proactive content blocking, visibility reduction, and removal (Allcott and Gentzkow 2017) (see Table 7).

Theme	Findings	Exemplary Sources
Facilitation – Many-to-Many Relationships		
Infrastructure provider	Social media platforms provide infrastructure enabling many-to-many relationships between opinion leaders and audiences. Fake News spreads via platform-specific mechanisms such as content sharing, commenting, and recommendation algorithms.	(Balmas 2014; Fourney et al. 2017; Guo and Zhang 2020; Rovetta and Bhagavathula 2020)
Facilitation – Enabling Platform Particularities		
Platform size	Smaller platforms (e.g., Gab), characterized by heterogeneous information and limited moderation, foster polarization and thus facilitate the dissemination of Fake News.	(Clark and Stoakes 2023; Etta et al. 2022; Peucker and Fisher 2022)
Media type	Platform media types (textual vs. visual) show limited impact on Fake News spread. However, platforms emphasizing interpersonal communication (e.g., WhatsApp, Telegram) notably facilitate Fake News dissemination.	(Neyazi et al. 2022; Park and Chai 2024; Valenzuela et al. 2022)
Mitigation – Platform Policy		
Moderation mechanisms	Platforms vary significantly in effectiveness at mitigating Fake News through proactive content blocking, visibility reduction, and removal strategies.	(Allcott and Gentzkow 2017; Etta et al. 2022; Peucker and Fisher 2022; Zannettou et al. 2017)

Table 7. Intermediary Characteristics Facilitating or Mitigating Fake News

Regulatory Environment

The regulatory environment defines the legal framework under which social media platforms regulate online speech and misinformation (Andorfer 2018). The literature identifies two primary challenges: first, determining whether platforms or governments should bear responsibility for mitigating Fake News; second, striking an appropriate balance between intervention efforts against Fake News and safeguarding fundamental human rights (e.g., freedom of expression, privacy protection, prevention of discrimination).

Advocates for government-led regulation argue that social media platforms function as critical infrastructures, similar to public utilities (e.g., power grids), and thus require stringent governmental oversight (Iosifidis and Andrews 2020). This perspective holds that governments have the duty to enforce strict regulations and hold users accountable for disseminating misinformation, explicitly excluding Fake News from freedom-of-expression protections.

Conversely, critics caution against strong governmental regulation, highlighting risks of infringing upon free speech and civil liberties. For example, while Wood and Ravel (2018) suggest that platforms should increase transparency regarding opinion leaders' social media dissemination, they caution against direct government content regulation, as it can undermine free expression. We have also observed the weaponization of government regulations, such as requests made to platforms to take down content criticizing a government (Kaye 2019).

The literature further highlights important trade-offs and proposes balanced regulatory measures. While criminalizing Fake News dissemination effectively reduces misinformation, such approaches inherently restrict freedom of expression (Helm and Nasu 2021b). Alternative regulatory approaches, such as enhancing media literacy and critical thinking among citizens, empower audiences to independently counter misinformation without compromising essential freedoms (Pang et al. 2022). Ultimately, effective regulation should rely on evidence-based policymaking, transparently communicated to build public trust and acceptance (Li et al. 2021) (see Table 8).

Theme	Findings	Exemplary Sources
Mitigation – Balance Between Censorship and Fake News Interventions		
Pro government-led regulations	Governments should strictly regulate content on social media, treating these platforms as critical infrastructures. Fake News dissemination should be explicitly excluded from freedom-of-expression protections, with individual users held accountable.	(Anansaringkarn and Neo 2021; Daud and Zulhuda 2020; Iosifidis and Andrews 2020; Rivas and Rausseo 2024; Torreblanca 2023)
Contra-government regulations	Government-imposed content regulations are criticized as infringements on freedom of expression, often leading to public backlash over perceived civil rights violations. Such regulations are also vulnerable to misuse, exemplified by cases of content removal, account blocking, or preemptive agenda-setting (e.g., Singapore, India).	(Grossman 2022; Li et al. 2021; Marecos et al. 2023; Slutskiy and Boonchutima 2022)
Tradeoffs and suggestions for regulatory measures	Criminalization of Fake News dissemination may effectively reduce misinformation but also restricts free speech. Regulations empowering audiences through media literacy and critical thinking initiatives can mitigate Fake News without compromising freedom of expression. Effective regulations should involve evidence-based policy decisions and transparent public communication to ensure broad support.	(Helm and Nasu 2021b; Li et al. 2021; Pang et al. 2022; Schlag 2023; Sievi and Pawelec 2025)

Table 8. Regulatory Environments: Balancing Freedom of Expression and Fake News Interventions

Communication Patterns and Fake News

X1…n Messages Containing Information About an Event

The Fake News context is uniquely characterized by fabricated information designed to mimic legitimate news sources, yet it circumvents rigorous journalistic processes (Lazer et al. 2018). Such fabricated messages frequently do not depend on real events, allowing malicious opinion leaders to entirely invent stories. Message fabrication occurs across a continuum, from fully fictional accounts of nonexistent events (Miletskiy et al. 2019) to messages that strategically integrate truthful elements, enhancing their perceived credibility (Mourao and Robertson 2019). The emergence and use of sophisticated technologies, such as deepfakes, further complicate authenticity verification, significantly strengthening

the deceptive effectiveness of fabricated content (Ahmed 2021; Himma-Kadakas and Ojamets 2022; Mourao and Robertson 2019) (see Table 9).

Theme	Findings	Exemplary Sources
Facilitation – Message Fabrication		
No reliance on actual events	Fake News frequently does not rely on real-world events but rather involves entirely fabricated messages crafted to appear authentic.	(Himma-Kadakas and Ojamets 2022; Lazer et al. 2018; Miletskiy et al. 2019)
Levels of fabrication	Fake News messages vary from fully fabricated stories about non-existent events to partially fabricated messages containing verifiable elements, strategically included to enhance perceived credibility. Advanced forms, such as deepfakes, significantly complicate authenticity verification efforts.	(Ahmed 2021; Himma-Kadakas and Ojamets 2022; Miletskiy et al. 2019; Mourao and Robertson 2019)

Table 9. Xn Messages: Facilitating Fake News Through Message Fabrication

X^{I} Publicized Online by Opinion Leaders

Malicious opinion leaders leverage unrestricted online platforms to swiftly disseminate large volumes of Fake News without the constraints of traditional journalistic standards (Larsson 2020; Lopez-Garcia et al. 2019; Miletskiy et al. 2019). They employ several specific strategies to enhance the effectiveness of misinformation.

These strategies include discrediting competing opinion leaders—particularly conventional news media—while positioning themselves as exclusive sources of truth (Das and Schroeder 2021; Ross and Rivers 2018). Another common tactic involves mimicking traditional news media formats and presentation styles, thus increasing the perceived authenticity and credibility of Fake News (Gurba et al. 2019). Malicious opinion leaders often appeal to negative emotions and sensationalism to increase audience engagement and susceptibility (Corbu et al. 2021; Santosa et al. 2018). Additionally, they frequently appeal to credible entities, instrumentalizing reputable experts, political figures, and mainstream media sources to bolster the perceived reliability of their claims (Canavilhas et al. 2019; Peucker and Fisher 2022) and aligning their messaging with dominant societal narratives (Bryanov et al. 2023). Furthermore, they employ strategies of reviving and repeating Fake News—resurfacing previously discredited information—to sustain its relevance and credibility (Corneille et al. 2020; Shin et al. 2018). Lastly, malicious opinion leaders systematically shield themselves from counterarguments by cultivating hostile environments that discourage or punish criticism and alternative interpretations (Nguyen 2020) (see Table 10).

Theme	Findings	Exemplary Sources
Facilitation – Opinion Leaders Leveraging the Online Information Ecosystem		
Liberated information dissemination	Malicious opinion leaders use social media to reach large audiences rapidly without adhering to journalistic standards (e.g., verification, ethical guidelines). This enables swift, high-volume dissemination of Fake News.	(Larsson 2020; Lopez-Garcia et al. 2019; Matherly and Greenwood 2024; Miletskiy et al. 2019)
Facilitation – Malicious Opinion Leader's Common Strategies		
Discrediting	Malicious opinion leaders actively undermine competing opinion leaders, positioning themselves as exclusive sources of truth.	(Bellutta et al. 2021; Das and Schroeder 2021; Kluknavská et al. 2025; Ross and Rivers 2018; Zimmermann and Kohring 2020)
Mimicking	Imitating legitimate news media formats and presentation styles enhances the perceived credibility of Fake News messages.	(Gurba et al. 2019; Janze and Risius 2017; Schneiders 2023; Yoshikawa et al. 2021)
Appeal to emotion	Exploiting audience vulnerabilities through emotional appeals, sensationalism, and "breaking news" framing increases Fake News impact.	(Calvillo and Harris 2022; Corbu et al. 2021; Illia et al. 2024; Santosa et al. 2018; Tantau et al. 2018)
Appeal to credible entities	Malicious opinion leaders enhance their own perceived credibility by referencing reputable third-party sources or prominent grievance narratives.	(Bryanov et al. 2023; Canavilhas et al. 2019; Lazar and Pop 2021; Peucker and Fisher 2022)
Revive and repeat	Repeated dissemination of Fake News, even after debunking, as well as reintroducing previously faded misinformation ("revival"), maintains its relevance and perceived credibility.	(Cohen et al. 2024; Corneille et al. 2020; Ng et al. 2021; Shin et al. 2018)
Shielding	Systematic rejection of alternative interpretations and fostering environments hostile to criticism (e.g., aggressive backlash) protect malicious narratives.	(Chibuwe 2020; Farkas 2023; Nguyen 2020)

Table 10. XI Malicious Opinion Leaders' Strategies and Exploitation of the Online Information Ecosystem

X[II] Intermediaries' Content Curation on Social Media Platforms

Social media content curation primarily occurs through algorithms determining what information audiences consume (Morales-i-Gras 2020). Platforms design these algorithms to maximize user engagement and attention, ultimately supporting their advertising revenue streams (Tufekci 2014; Zuckerberg 2018). As a result, these algorithms can inadvertently amplify confirmation bias, facilitating the spread of Fake News by repeatedly exposing users to content similar to their previous interactions (Balmas 2014).

To mitigate such inadvertent amplification, platforms use automated detection and flagging systems (Losifidis and Nicoli 2020). They also introduce measures to enhance the visibility of fact-checking, such as displaying corrective information prominently alongside false content (Van Heekeren 2020). However, evidence regarding the effectiveness of flagging is mixed. Some studies suggest that flagging Fake News can unintentionally increase its virality (Greene and Murphy 2021), particularly if users perceive the flags as driven by political or ideological biases (Lee et al. 2023). Conversely, flags generated by neutral, AI-driven models or community-based crowdsourcing show higher effectiveness in reducing Fake News believability (Chung et al. 2023). In comparison, crowdsourced flagging opportunities or flags by artificial intelligence (AI) show the highest efficacy in persuading audiences (Chung et al. 2023). Finally, platforms also employ inoculation strategies, also known as "pre-bunking," proactively exposing users to warnings and minor doses of misinformation tactics to strengthen their resistance to Fake News (Van der Linden et al. 2017) (see Table 11).

Theme	Findings	Exemplary Sources
Facilitation – Platform Distribution Mechanisms		
Algorithmically enforced confirmation bias	Platform algorithms prioritize and distribute content most likely to engage users (e.g., likes, comments, shares). Consequently, these algorithms amplify Fake News that aligns with users' existing beliefs, exacerbating confirmation bias.	(Balmas 2014; Morales-i-Gras 2020; Schinello 2024; Tufekci 2014; Zuckerberg 2018)
Mitigation – Anti-Fake News Initiatives		
Fact-checking and flagging	Platforms utilize automated and AI-driven models to detect and flag Fake News. Additionally, audiences participate in crowdsourced flagging or contribute contextual information (e.g., community notes on platforms like Twitter/X).	(Dobber et al. 2023; Etta et al. 2022; Losifidis and Nicoli 2020; Moravec et al. 2023; Sheen 2024)
Flagging efficacy	The effectiveness of Fake News flags significantly varies with their design. Flags perceived as politically or ideologically biased have limited effectiveness and can inadvertently amplify the flagged content. Conversely, flags perceived as community-driven or neutral (e.g., AI-generated) exhibit high efficacy in mitigating Fake News.	(Chung et al. 2023; Greene and Murphy 2021; Lee et al. 2023)
Inoculation	Providing general warnings about tactics employed by malicious actors, and introducing minor exposure to Fake News as part of inoculation efforts, significantly reduces its believability.	(Lees et al. 2023; Van der Linden et al. 2017; van Der Linden et al. 2020; Yu and Yan 2024)

Table 11. Content Curation Algorithms and Their Impact on Fake News Dissemination and Mitigation

IO Intermediary Moderates Opinion Leader

The literature suggests two primary approaches through which platforms moderate opinion leaders at the account level. First, intermediaries may directly ban individual opinion leader accounts to halt the spread of Fake News (Bak-Coleman et al. 2022). Second, intermediaries utilize visibility restrictions, such as shadow banning, to reduce the audience reach of problematic opinion leaders without removing

their accounts entirely (Srivastava et al. 2019b). Despite these measures, malicious opinion leaders frequently create alternate accounts to bypass bans (please see section "opinion leaders"), thereby limiting the overall effectiveness of account removal as a Fake News mitigation strategy (Linvill and Warren 2020; Srivastava et al. 2019b) (see Table 12).

Theme	Findings	Exemplary Sources
Mitigation – Platforms' Account Moderation Practices		
Forms of account moderation	Platforms moderate opinion leaders by removing individual accounts or restricting their reach through reduced visibility (e.g., shadow banning).	(Bak-Coleman et al. 2022; Fierens and Rêgo 2024; Papanastasiou 2020; Srivastava et al. 2019b)
Removal effectiveness	Account removal demonstrates limited effectiveness, as malicious opinion leaders quickly circumvent bans by creating alternate accounts to sustain Fake News dissemination.	(Bak-Coleman et al. 2022; Helm and Nasu 2021a; Prochaska et al. 2023; Srivastava et al. 2019a)

Table 12. Intermediaries' Moderation of Opinion Leaders via Account Removal or Restrictions

AI Audience Directly Addresses Intermediaries

Audience members directly report Fake News to intermediaries, though existing research remains limited regarding their motivations for reporting (Coscia and Rossi 2020). A general decline in trust toward social media has increased adversarial reporting, wherein audiences label legitimate news from traditional media as Fake News without verification (Luo et al. 2022; Valdez and Ziefle 2018). Social context also shapes reporting behavior; audiences are less inclined to report Fake News if it aligns with their community's norms or if strong social ties discourage scrutiny (Aljasir and Bjanaid 2021). Conversely, content that breaches community norms or popular opinion is more frequently reported (Gimpel et al. 2021). To enhance audience-driven Fake News reporting, platforms should prioritize usability, implementing streamlined systems that enable reporting with minimal effort (Zhou et al. 2023).

Theme	Findings	Exemplary Sources
Facilitation – Lack of Community Reporting		
Adversarial reporting	Audiences increasingly report legitimate news media content as Fake News due to a generalized decline in trust within social media environments.	(Luo et al. 2022; Valdez and Ziefle 2018; Zhou et al. 2023)
Social context	Strong social ties with opinion leaders or other audience members decrease individuals' willingness to verify or report Fake News. Audience members in closed or distinct social media communities (e.g., Facebook groups, subreddits) are especially unlikely to report misinformation circulating within these groups.	(Aljasir and Bjanaid 2021; Gimpel et al. 2021; Waruwu et al. 2021)
Mitigation - Audience Support for Fake News Reporting		
Reporting usability	User-friendly reporting mechanisms requiring minimal effort (e.g., only a few clicks) enhance audience engagement in reporting Fake News.	(Zhou et al. 2023)

Table 13. Audience-Initiated Fake News Reporting to Intermediaries

AO Audience Engages with Opinion Leaders via Intermediaries

Audiences facilitate the spread of Fake News by selectively curating information based on their trust levels and confirmation biases toward opinion leaders (Kaiser et al. 2022). As trust in opinion leaders declines due to pervasive misinformation (Park et al. 2020), audience members increasingly follow opinion leaders who reinforce their pre-existing beliefs, inadvertently making themselves more susceptible to Fake News (Nekmat 2020). For example, during elections, audiences tend to follow politically aligned opinion leaders, selectively amplifying partisan misinformation that aligns with their ideological views.

Conversely, audience members mitigate Fake News dissemination by actively managing their interactions with opinion leaders. For example, audiences collectively engage in content verification, critically assessing and challenging opinion leaders' posts (Corsi et al., 2024; Tully and Singer, 2023). If audiences identify opinion leaders spreading false information, they may unfollow or block these individuals, thus reducing their visibility, reach, and ultimately diminishing their social capital and influence (Guo et al. 2022). In cases where opinion leaders are found disseminating misinformation, audience members engage in shaming (Babcock et al., 2019), which can escalate into online "firestorms"—rapid, collective negative reactions that significantly diminish opinion leaders' credibility (Jang et al. 2019).

Theme	Findings	Exemplary Sources
Facilitation– Content Curation Through Following		
Self-reinforced confirmation bias	Audiences actively curate content through social media functions such as following opinion leaders, thus reinforcing confirmation biases. Following malicious opinion leaders directly facilitates exposure to Fake News.	(Kaiser et al. 2022; Nekmat 2020; Park et al. 2020)
Mitigation – Reaction Toward Identified Malicious Opinion Leaders		
Blocking of malicious opinion leaders	Audiences mitigate Fake News by blocking identified malicious opinion leaders, reducing their visibility and reach.	(Bakshy et al. 2015; Hartley and Khuong 2020; Kaiser et al. 2022; Mourão and Molyneux 2021)
Audience response to opinion leaders sharing Fake News	Audiences respond to malicious opinion leaders by engaging in direct interactions, such as sending corrective messages, shaming them, or initiating online "firestorms." Audiences individually or collectively verify opinion leaders' shared content to manage misinformation.	(Babcock et al. 2019; Corsi et al. 2024; Jang et al. 2019; Paisana et al. 2020; Tully and Singer 2023)

Table 14. Audiences Engage with Opinion Leaders via Active Curation through Following, Blocking, and Direct Interactions

AA Message Redistribution Among Audiences

Fake News spreads through social media engagement features, such as liking, sharing, and commenting (Buchanan and Benson 2019). As such, audiences play a central role in disseminating not only authentic content but also misinformation (Castioni et al. 2022). The literature identifies several audience-driven motivations behind Fake News sharing. For instance, social influence significantly motivates sharing behaviors, as individuals distribute Fake News to initiate conversations or maintain social ties, especially with those they value (Hadlington et al. 2023; Jost et al. 2020). Status-seeking behavior also encourages sharing Fake News quickly to position oneself as an influential informant, thus boosting social recognition and encouraging further sharing (Park and Lee 2020). Additionally, altruism motivates audiences to distribute misinformation under the belief that they are genuinely helping inform others (Apuke and Omar 2021). Hedonistic reasons, such as seeking entertainment, combating boredom, or the thrill of rapidly sharing new information, also drive the spread of Fake News (Balakrishnan et al. 2021; Zhang et al. 2022b).

Fake News campaigns exploit these audience behaviors to amplify misinformation. Historically, these campaigns largely depended on bots to disseminate Fake News (Linvill et al. 2019). Recently, however, there has been a notable shift from purely bot-driven dissemination to participatory campaigns involving real users actively sharing misinformation (Starbird 2019; Wilson and Starbird 2021). In these participatory campaigns, genuine human users unintentionally reinforce false narratives within online communities. Such dynamics contribute to constructing alternative realities and can escalate to significant real-world mobilization, exemplified by events such as the January 6th Capitol riots (Nguyen 2020; Prochaska et al. 2023).

Theme	Findings	Exemplary Sources
Facilitation – Factors that Trigger Fake News Sharing		
Social influence	Audience members are influenced to share Fake News when they observe supportive comments or high engagement (likes, shares, comments) from other users. Sharing is often motivated by desires to socialize, initiate conversations, or strengthen relationships.	(Castioni et al. 2022; Colliander 2019; Hadlington et al. 2023; Jost et al. 2020; Zhang et al. 2024)
Status seeking	Audience members quickly share content, including Fake News, to position themselves as key informants or influential figures within their networks, thereby encouraging future engagement and reinforcing their perceived social status.	(Carrera 2023; Chaudhuri et al. 2025; Malik et al. 2023; Park and Lee 2020)
Altruism	Audience members share Fake News with genuine intentions to inform others, unaware of the misinformation, believing that such content will benefit their social networks.	(Apuke and Omar 2021; Carrera 2023; London Jr et al. 2022; Pennycook et al. 2020b; Umejei et al. 2024)
Hedonism	Audience members share Fake News out of a desire for gratification, entertainment, boredom relief, or the excitement of being the first to spread novel information.	(Balakrishnan et al. 2021; Mansoori et al. 2023; Mazhar et al. 2024; Zhang et al. 2022b)

Facilitation – Weaponization of Many-to-Many Communication		
Bots redistribute Fake News	Bots significantly amplify Fake News by repeatedly sharing and mimicking genuine audience behaviors, effectively manipulating and distorting public discourse.	(Himelein-Wachowiak et al. 2021; Linvill et al. 2019; Ross et al. 2019; San Martin et al. 2020)
Weaponization of audiences	Participatory disinformation campaigns deliberately seed Fake News within targeted communities, exploiting social media interactions (shares, comments, endorsements). Genuine audience members unknowingly redistribute Fake News, unintentionally contributing to the construction of alternate realities and distorted narratives.	(Ippa et al. 2024; Nguyen 2020; Prochaska et al. 2023; Starbird 2019; Wilson and Starbird 2021)

Table 15. AA Results on Participatory Disinformation Campaigns and Fake News Redistribution Facilitators

X^{III} Intermediaries Filter Original Content

Beyond account and content moderation, intermediaries proactively moderate uploads using filters as alternative methods to mitigate Fake News (Llansó 2020). This proactive approach is typically driven by regulatory requirements (see Regulatory Framework). Platforms employ hash-matching techniques, comparing newly uploaded content to digital fingerprints stored in databases of restricted materials. Such methods effectively detect and block prohibited content, including repeated uploads involving copyright infringement (Youtube 2019), or images related to child sexual abuse (Microsoft 2019).

Theme	Findings	Exemplary Sources
Mitigation – Proactive Moderation		
Keyword filtering	Scans post content for specific words commonly found in Fake News, designated as prohibited by regulators.	(Llansó 2020; Perel and Elkin-Koren 2022)
Hash-matching	Employs unique digital fingerprints (hash values) to identify and block Fake News content by comparing these values against databases of restricted material.	(Llansó 2020; Microsoft 2019; Perel and Elkin-Koren 2022; Youtube 2019)

Table 16. Intermediaries Proactively Moderate Content Before Posting on Social Media Platforms Based on Keyword Filters or Hash-Matching

Discussion

This study introduces the Model of Online Networked Communication (MONC), a theoretical framework explicitly designed to capture how social media platforms have fundamentally transformed traditional communication processes, creating new pathways for phenomena such as Fake News dissemination. Employing a grounded theory-based literature review method (Wolfswinkel et al. 2013), we systematically synthesized a multidisciplinary body of research on Fake News, culminating in the development of MONC. Specifically, MONC delineates five clearly defined communication roles—Events, Opinion Leaders, Intermediaries, Audience, and Regulatory Environment—and maps eight

distinctive interaction pathways (X1…n, XI, XII, IO, AI, AO, AA, and XIII), thus providing a structured analytic lens for comprehending contemporary digital communication complexities.

Our findings reveal several critical characteristics unique to social media environments, such as algorithmic content curation, dynamic and influential roles of opinion leaders, audience-driven content redistribution, and explicit regulatory interventions, all of which collectively shape the dissemination and mitigation of Fake News. These findings extend traditional linear communication models, notably Westley and MacLean's (1957), by explicitly integrating multidirectional interactions and automated gatekeeping mechanisms. For instance, audience behaviors such as selective content sharing, direct platform engagement (e.g., reporting misinformation), and collective responses (e.g., shaming of misinformation spreaders) highlight the audience's active participation in gatekeeping roles. Empirical cases, including COVID-19 misinformation campaigns and election-related polarization, concretely illustrate these complex dynamics, grounding MONC's theoretical insights firmly within real-world contexts.

The subsequent sections detail the specific contributions of each identified communication role and pattern, providing novel theoretical insights and practical implications closely aligned with our research objectives.

Future Research Directions

While the MONC effectively summarizes state-of-the-art research on factors facilitating and mitigating Fake News, it also provides a structured framework to guide future research directions. Our analysis indicates that (1) existing research predominantly addresses events as contextual drivers (E) and audience characteristics (A), as well as opinion leaders' dissemination strategies (XI) and audience-driven message redistribution (AA). Consequently, we explicitly call for research addressing prominent knowledge gaps concerning (2) specific communication role characteristics (notably opinion leaders, intermediaries, and regulatory environments), (3) processes of information creation (X1…n), (4) mechanisms of content moderation (IO, AI, AO), and (5) platform filtering techniques (XII, XIII). Accordingly, the subsequent section provides detailed future research recommendations for each identified role and communication pattern ((1)–(5)), highlighting critical opportunities for advancing our theoretical and practical understanding of Fake News dynamics within the online information ecosystem (see Table 17).

While an extensive body of literature addresses the roles of events (E) and the audience (A), further research could provide deeper theoretical and empirical insights into their influence on Fake News dissemination. Currently, events predominantly serve as contexts rather than focal points in Fake News studies (e.g., Gallotti et al. 2020; Guess et al. 2020). Consequently, the specific event characteristics that might facilitate the emergence of Fake News remain largely unexplored. Future research should explicitly examine how particular event attributes interact with audience biases, clarifying when and why certain events become precursors to Fake News campaigns.

Additionally, different event types produce distinct Fake News narratives, illustrated by cases such as the train accident in East Palestine (Tolentino and Sung 2023). For example, public health crises like COVID-19 triggered widespread misinformation regarding vaccine efficacy, whereas natural disasters often prompt localized rumors about governmental aid and response effectiveness (see Table 4). Given this variability, future studies should systematically investigate appropriate mitigation strategies tailored explicitly to the type of event, examining how specific characteristics might necessitate different response approaches to effectively manage misinformation impacts.

The audience role also warrants further research attention, particularly given its heterogeneity. Audiences comprise diverse subgroups with varying ideologies, backgrounds, and susceptibilities to Fake News. Current research insufficiently explores the characteristics and roles of these specific subgroups in facilitating misinformation spread. For instance, certain demographic or social groups may experience disproportionately negative economic or social consequences from Fake News (Cousins 2020). Similarly, politically affiliated subgroups, such as partisan communities on social media platforms, might disproportionately amplify politically charged misinformation, leading to intensified polarization and reduced trust in democratic institutions (Kitchens et al. 2020).

Moreover, the literature increasingly highlights the audience's critical role in mitigating misinformation. Thus, we propose expanding research on the concept of “desistance”, examining circumstances under which individuals or groups actively choose to disengage from Fake News (Farrington 2007). Such research could also investigate how users transition from passive disengagement to active countering behaviors, such as publicly correcting misinformation or educating peers. Furthermore, exploring factors that transform initially passive users into proactive "digital gatekeepers" would offer valuable insights for designing platform-based or educational interventions. Such work would significantly enhance our understanding of audience agency and resilience within online communication ecosystems.

Comparably, the extensive body of research on audience redistribution and opinion leadership can be further expanded. Various forms of Fake News often originate from fringe platforms and then spread onto mainstream platforms through the mechanism of audience redistribution (AA). However, we currently lack an understanding of the processes through which audiences actively cultivate, reframe, or legitimize Fake News content, thereby enabling its seamless transfer and broader acceptance on mainstream social media platforms. With the emergence of participatory disinformation campaigns (Starbird 2019; Wilson and Starbird 2021), we propose further investigations into Fake News as a cultural phenomenon (Nguyen 2020; Prochaska et al. 2023), specifically examining how cultural narratives, symbols, or values contribute to the normalization and integration of misinformation within online communities.

Furthermore, we should explore how platforms can actively incentivize audiences to avoid sharing Fake News. While malicious opinion leaders consistently develop effective strategies to disseminate Fake News (XI), traditional news media organizations often struggle to effectively counter these narratives or strategically reposition their messaging. For example, during the 2020 U.S. presidential election, traditional media outlets faced significant challenges in swiftly counteracting misinformation about

mail-in voting fraud propagated by influential partisan figures (Benkler et al. 2020). To effectively counteract malicious opinion leadership, we call for research into specific strategies journalistic institutions could adopt to proactively or reactively debunk misinformation, enhance narrative credibility, and limit the impact of Fake News after its dissemination.

The identity of a messenger can significantly shape how audiences interpret and trust misinformation. Sterrett et al. (2019) demonstrated that messengers often influence audience trust in misinformation more significantly than the content itself. For example, misinformation about health risks may be more readily accepted when shared by popular celebrities or trusted community figures rather than anonymous online accounts (Hoffman and Tan 2013). However, current literature provides limited insights into how audiences evaluate credibility differently when misinformation is conveyed by varying types of opinion leaders or peers. Thus, further research should specifically examine the mechanisms through which messenger characteristics—such as perceived authority, trustworthiness, or social identity—influence audience perceptions of misinformation credibility and acceptance.

Role/Pattern	Research Objective	Exemplary Research Questions
Stream 1: How does the online information ecosystem facilitate Fake News?		
Events	Overall relevance of event characteristics	What role do events play compared with existing biases in deceiving audiences?
Opinion Leader	Determinants for popular opinion leaders	What factors corrupt opinion leaders to become malicious opinion leaders?
Audience	Target group specificity	How do distinct communities jointly facilitate Fake News? What is the role of marginalized communities in the creation and spread of Fake News?
Intermediary	Platform differences	How does the lack of moderation facilitate the creation and spread of Fake news? How does the technocentric approach of platforms facilitate the creation and spread of Fake News?
Regulatory environment	Governmental interference	How do platforms navigate governmental pressure to transparently disclose content moderation decisions?
$X_{1...n}$	Fabrication of Fake News through AI	How does the use of generative AI for sourcing information facilitate Fake News production?
X^{I}	Messenger effect	How do the characteristics of social media accounts moderate the effect of confirmation bias on news believability?
X^{II}	Adverse effects and tradeoffs of content moderation	What are the tradeoffs (e.g., trust, privacy, polarization, and well-being) between existing Fake News content moderation mechanisms?
IO	Moderation of popular opinion leaders	What are the ramifications (e.g., Streisand-effect: increased redistribution of Fake News by target audiences) of popular forms of opinion leader moderation, such as banning and shadow-banning?
AI	Reporting weaponization	How are reporting systems weaponized?
AO	Following behaviors	What determines an audience member's behavior to follow malicious opinion leaders? What are the differential effects of algorithmic audiencing versus opinion leadership on exposure to Fake News?

AA	Redistribution across platforms	How do audiences redistribute Fake News across different platforms? How do changes in information distribution contribute to the emergence of Fake News as a cultural phenomenon?
X^{III}	Ramifications of upload filters	How do upload filters affect trust in platforms, governments, and news media?
Stream 2: How can the online information ecosystem successfully mitigate Fake News?		
Events	Anticipation of Fake News event triggers	Which response strategies help mitigate disinformation campaigns triggered by particular events?
Opinion Leader	Ideological opposition to Fake News	How can news media organizations establish trust amid a Fake News environment?
Audience	Positive transformation of audiences	What drives audience desistance from Fake News?
Intermediary	Responsible platform guidelines	How can intermediaries design ethical moderation policies?
Regulatory environment	Government-led interventions	How can governments impose transparency requirements for Fake News moderation on certain platforms?
$X_{1...n}$	Establish source authenticity	How do we detect and communicate message reliability (e.g., watermark)?
X^{I}	Opinion leaders' response strategies	How can opinion leaders ensure the timely and pervasive debunking of Fake News?
X^{II}	Improving content moderation practices	How should we design content curation algorithms that optimize long term engagement? What are new forms of content moderation that are suitable for mitigating Fake News?
IO	Curating constructive content creation	How can platforms incentivize desirable opinion leader behavior?
AI	Audience empowerment	What drives audiences to report Fake News?
AO	Unfollowing and blocking antecedents	When do audiences unfollow malicious opinion leaders? What techniques do grassroots movements deploy to counter Fake News?
AA	Cultivating positive platform engagement	Which platform-based interventions discourage the redistribution of Fake News (e.g., AI-generated prompts)?
X^{III}	Pro-active content moderation	How should upload filters be designed to minimize negative ramifications? What is the efficacy of upload filters in preventing Fake News?

Table 17. Future Research Directions on the Facilitation and Mitigation of Fake News with Objectives and Research Questions

Expanding beyond research gaps, we turn to identifying broader and more fundamental areas requiring deeper theoretical and empirical exploration—specifically, the nuanced characteristics of opinion leaders, intermediaries, and regulatory environments. During societal emergencies, certain opinion leaders become highly vocal, often creating or reproducing conspiratorial narratives (Dow et al. 2021). However, it remains unclear precisely which factors or conditions trigger opinion leaders' transition into malicious actors, particularly in crisis situations. For instance, the COVID-19 pandemic saw previously credible public figures and influencers rapidly shifting towards disseminating conspiracy theories and misinformation about vaccine safety. Consequently, we need deeper insights into the mechanisms of

radicalization, examining how opinion leaders transform and eventually become active purveyors of Fake News.

Addressing the erosion of audience trust and transparency within algorithmically mediated ecosystems is crucial for combating Fake News. Research has established a general loss of trust in news media (Valdez and Ziefle 2018). For example, during the COVID-19 pandemic, inconsistent messaging from mainstream media and public authorities about mask efficacy contributed significantly to public confusion and diminished trust (Ho and Huang 2021; Sauer et al. 2021). Given this context, it is crucial for researchers to systematically investigate strategies through which opinion leaders—particularly credible news media organizations and journalists—can effectively rebuild audience trust within the increasingly polarized and misinformation-prone online information ecosystem.

As intermediaries become more reliant on algorithmic content moderation, their decision-making processes have become increasingly opaque (Katzenbach 2021; Shaban 2020). Algorithms, by heavily influencing users' media exposure, reinforce and amplify existing biases (Balmas 2014). For instance, YouTube's recommendation algorithm has been criticized for guiding users towards increasingly polarized and conspiratorial content, inadvertently amplifying misinformation (Yesilada and Lewandowsky 2022). Consequently, it is essential to further examine how such technocentric approaches to content moderation might inadvertently exacerbate issues like algorithmic bias, marginalization of minority viewpoints, and systematic misclassification of legitimate content—thus unintentionally facilitating, rather than mitigating, the spread of Fake News.

Meanwhile, by leveraging the regulatory environment, governments may exert significant pressure on intermediaries, including threats of legal repercussions, to compel the removal of politically unfavorable content (Harbath 2023). For example, recent demands from governments in countries such as India and Turkey have pressured platforms like Twitter and Facebook to remove politically sensitive content or face potential restrictions and legal consequences. Thus, we need to better understand how platforms navigate and respond to such government pressures, particularly those posing potential legal threats to platform employees. Further research could provide essential guidance for intermediaries, helping them adopt ethical countermeasures aligned with the "safety by design" framework. Similarly, we advocate for research that explores how regulatory bodies can implement clearer, more transparent communication practices in their moderation requests, helping to prevent problematic practices such as jawboning (Bambauer 2015).

Considering the significant knowledge gaps surrounding information creation, it is critical to examine the implications of news organizations increasingly utilizing generative AI to source and interpret vast amounts of information (X1…n) (Karanasios and Risius 2024; Stenberg 2024). For instance, automated AI-generated news summaries may inadvertently integrate misinformation from unreliable sources, thereby reducing overall transparency and complicating information verification. With the expanding reliance on AI in content creation, it becomes progressively challenging to determine information reliability. Consequently, future research should explicitly investigate how AI-driven sourcing practices might unintentionally facilitate the production and dissemination of Fake News. Additionally, as deepfakes significantly complicate judgments about message authenticity (Ahmed 2021; Himma-

Kadakas and Ojamets 2022), we propose that researchers actively explore and develop effective methods—such as digital watermarking techniques—to clearly differentiate authentic content from AI-generated or otherwise fabricated media.

Optimizing content filtering processes demands careful navigation of complex trade-offs that may unintentionally amplify misinformation. Current approaches to content moderation (XII) inevitably involve difficult trade-offs (Jiang et al. 2023), such as the implied truth effect of Fake News flagging (Lee et al. 2023; Pennycook et al. 2020a). For example, labeling certain content as misleading can inadvertently strengthen users' beliefs in unflagged, yet potentially unreliable, information. Given such trade-offs, further research should systematically examine the content moderation remedies outlined by Goldman (2021), specifically comparing the effectiveness, unintended consequences, and broader impacts of various moderation strategies.

Content moderation practices often prioritize highly visible, provocative posts that drive immediate reactions (Tufekci 2014), inadvertently fueling the spread of Fake News and undermining long-term audience engagement. For instance, content that provokes strong emotional reactions, such as outrage or fear, frequently gains rapid attention and short-term visibility. However, as platforms become known for hosting such provocative content, audiences may ultimately lose trust and disengage. To address this, future research should investigate alternative content curation strategies that emphasize sustainable, long-term user engagement and retention rather than immediate attention.

Platforms have begun employing proactive moderation strategies (XIII) through upload filters (Llansó 2020), yet technologically sophisticated AI-mediated moderation methods remain relatively underexplored. These automated moderation methods raise significant concerns regarding transparency and potential infringement of civil liberties, particularly freedom of expression. For instance, automated filters on platforms like YouTube have occasionally removed legitimate journalistic reports or educational content due to misclassification, raising concerns over censorship and freedom of expression (Wired 2018). Thus, further research is necessary to understand how reduced transparency in these AI-driven upload filters might inadvertently undermine trust in intermediaries and facilitate, rather than mitigate, Fake News.

Moreover, the implementation of automated upload filters carries potential unintended consequences and complex trade-offs. For example, overly aggressive filters intended to block extremist content have unintentionally suppressed important discussions around political activism or historical documentation (Earl et al. 2022), highlighting the delicate balance platforms must maintain between moderation and open discourse. Consequently, researchers should systematically examine these trade-offs and develop guidelines to minimize negative side effects, ensuring a balanced approach between content moderation effectiveness and preserving civil liberties (Goldman 2021).

Finally, there is a significant lack of insight into the broader implications and effectiveness of current content moderation practices, highlighting an urgent need for deeper empirical and theoretical understanding. The emergence and persistence of Fake News illustrate that factual accuracy alone no

longer ensures audience attention within today's online information ecosystem. For instance, misinformation around climate change often gains higher initial engagement on social media platforms compared to factually accurate scientific reports, due to emotionally charged content that triggers user interaction (Treen et al. 2020). Therefore, we propose exploring innovative moderation mechanisms through which intermediaries can actively incentivize opinion leaders to refrain from disseminating misinformation—for example, by rewarding desirable platform behaviors through special privileges or enhanced visibility (IO).

Additionally, account-level moderation strategies, such as deplatforming opinion leaders, require closer examination, given potential unintended consequences like radicalization, shifts in communication styles, or migration to less-regulated platforms. For example, after prominent figures were banned from platforms like Twitter, they frequently migrated to alternative platforms such as Parler and Telegram, intensifying extremist discourse and reducing overall transparency (Ali et al. 2021). The DIME (disidentification, innovation, moralization, and energization) model (Louis et al. 2020) offers a valuable theoretical lens to systematically assess and understand these potential adverse outcomes, thereby informing more effective and nuanced moderation policies.

Audiences can also exploit Fake News reporting systems (e.g., AI-driven moderation tools) to silence and censor opinion leaders by triggering shadowbans (Nicholas 2022). Thus, future research must actively investigate methods to prevent the weaponization of reporting systems by audiences, ensuring that these mechanisms serve their intended purpose without stifling legitimate public discourse. For example, organized campaigns have used mass reporting to trigger automatic bans against political opponents, undermining free speech and open debate (Jhaver et al. 2019).

The same logic applies to audiences, who require stronger incentives to actively seek out accurate, fact-based reporting. Given that audiences play a critical role in reporting problematic online behaviors (Wong et al. 2021), further research should identify strategies to promote a positive online information ecosystem more resilient to Fake News. While previous research confirms that individuals receive a significant proportion of Fake News by following malicious opinion leaders (AO) (Bakshy et al. 2015; Robertson et al. 2023), we still lack insights into whether this occurs due to the inherent popularity of opinion leaders or through algorithm-driven recommendations (algorithmic audiencing) (Riemer and Peter 2021).

Furthermore, considering the online ecosystem enables direct interactions between audiences and opinion leaders—often manifesting as online firestorms or organized backlash (Jang et al. 2019; Starbird 2019; Wilson and Starbird 2021), — we call for additional research into the varied techniques grassroots movements employ to counteract Fake News, including controversial practices such as porn-bombing or coordinated online firestorms. For example, in response to misinformation campaigns on social media platforms, activist groups have coordinated mass-commenting tactics or hashtag hijacking to disrupt and counteract misleading narratives (Weber et al. 2022).

Contributions to Theory and Practice

This article set out to answer one research question: How do the characteristics of the online information ecosystem facilitate and mitigate Fake News? To address this question, we applied the grounded theory method to comprehensively review multidisciplinary literature consisting of 665 Fake News articles (Wolfswinkel et al. 2013). From this extensive review, we derived the Model of Online Networked Communication (MONC), explicitly designed to address critical limitations of traditional communication models such as Westley and MacLean's (1957), which inadequately account for the complex dynamics introduced by digital and social media platforms.

MONC contributes distinctively to existing literature by uniquely integrating five communication roles (Events, Opinion Leaders, Intermediaries, Audience, Regulatory Environment) and eight communication patterns, thereby explicitly capturing the multidirectional and algorithmically mediated interactions characteristic of contemporary social media ecosystems. Specifically, MONC clarifies significant theoretical gaps by addressing digitally induced shifts, including the replacement of traditional news media organizations by algorithmically curated social media platforms as intermediaries (Tandoc Jr and Vos 2016), the active and direct engagement between opinion leaders and audiences (Lewis and Westlund 2015), increasingly complex audience-driven redistribution of messages (Bro and Wallberg 2014), platform-driven algorithmic content curation (Tufekci 2014), reliance on audience cognitive vulnerabilities in processing Fake News (Pennycook and Rand 2019b), and inadequacies in existing regulatory frameworks (Helm and Nasu 2021b).

First, MONC explicitly updates and extends traditional communication frameworks, prominently Westley and MacLean's linear mass communication model, commonly referenced by IS scholars (George et al. 2018). While traditional models emphasize a linear flow from communicators through intermediaries to passive audiences, MONC highlights critical divergences such as active audience participation, the transformation of gatekeepers into opinion leaders, and algorithmically mediated interactions. Future research might empirically explore how these digitally mediated interactions specifically alter audience behavior and influence opinion leader strategies.

Second, MONC explicitly integrates and extends insights from a diverse array of disciplinary perspectives, including communication studies, information systems, psychology, political science, and platform governance. By synthesizing these multidisciplinary insights, MONC uniquely provides researchers with a comprehensive framework capable of addressing complex questions related to Fake News. The synthesis identified pivotal research themes such as strategies used by opinion leaders to promote or counter misinformation, event characteristics influencing Fake News dissemination, and the vulnerabilities and protective factors affecting different audience segments. Future studies could, for instance, investigate cognitive interventions aimed at reducing audience susceptibility to misinformation, leveraging psychological insights (Pennycook and Rand 2019b).

Third, MONC clearly specifies the mechanisms and outcomes crucial to understanding Fake News dynamics. Mechanisms articulated by MONC include algorithmic prioritization of engaging content by

digital platforms (Tufekci 2014), active redistribution of misinformation by engaged audiences (Chung 2023), and opinion leader strategies to disseminate targeted misinformation. These mechanisms explicitly lead to measurable outcomes such as enhanced spread and virality of Fake News, increased political polarization, and reduced trust in intermediaries and traditional media sources. Future research might evaluate platform-based interventions designed to mitigate these outcomes, such as content labeling or enhanced algorithmic transparency measures.

Through these explicitly articulated contributions, MONC substantially advances theoretical understandings of Fake News within digital ecosystems. Additionally, it identifies critical research gaps and offers structured research questions across MONC's dimensions, guiding future scholarship toward comprehensive investigations of intermediaries, opinion leaders, audiences, and regulatory frameworks.

Limitations

The contributions of this paper need to be considered in light of several limitations, each of which provides important directions for future research. First, our study relies on a systematic literature review, which, while thorough and comprehensive, inherently lacks empirical validation. This methodological constraint limits our ability to empirically verify the practical applicability and efficacy of MONC's constructs and mechanisms. Future research should therefore conduct empirical studies, such as surveys or experiments, to validate and refine MONC's roles and communication patterns in real-world settings.

Second, the applicability and generalizability of MONC might vary across different types of social media platforms or cultural contexts. Social media platforms differ significantly in their user base, content moderation policies, algorithmic design, and interaction affordances, potentially influencing Fake News dissemination dynamics differently. Additionally, cultural and political contexts may profoundly impact the behaviors of intermediaries, opinion leaders, and audiences. Future empirical research should systematically assess MONC's validity and generalizability across various platforms (e.g., Facebook, Twitter, TikTok) and diverse cultural contexts.

Third, operationalizing key constructs within MONC for empirical research presents significant challenges. Constructs such as algorithmic mediation, opinion leader influence, audience vulnerabilities, and regulatory effectiveness are conceptually nuanced and complex to measure empirically. Future studies need to develop clear operational definitions and robust measurement instruments for these constructs. For instance, research could employ machine learning techniques or computational analyses to quantify algorithmic mediation effects, use network analysis to assess opinion leader influence, or leverage experimental designs to test regulatory impacts empirically.

Finally, our communication-centered perspective on Fake News may have excluded alternative theoretical perspectives and relevant related topics such as misinformation, false news, rumors, and propaganda. For example, "propaganda" refers to a distinct form of information, often government-controlled and highly manipulative, though not necessarily false (Moravec et al. 2023). Therefore, we

encourage future research to integrate additional theoretical perspectives and explore complementary concepts to address the complexities inherent in online information ecosystems comprehensively.

Similarly, while our analysis systematically reviews the Fake News literature, additional related phenomena such as conspiracy theory radicalization (Abdalla Mikhaeil and Baskerville 2024), information consumption (e.g., filter bubbles and echo chambers; (Bozdag and van den Hoven 2015; Kitchens et al. 2020; Risius et al. 2019), or online extremism (Risius et al. 2023) could provide richer contexts for validating and expanding MONC. Although preliminary checks indicated no missed articles relevant to our specific scope, future research should extend the literature search by including additional keywords such as "disinformation" and triangulate findings with those from related research streams to enhance MONC's robustness and comprehensiveness.

Conclusion

Analyzing the extensive, interdisciplinary Fake News literature, we developed the MONC to succinctly capture prevailing communication roles and patterns characterizing contemporary online information ecosystems. By significantly extending Westley and MacLean's (1957) traditional model of mass communication through a grounded theory approach (Wolfswinkel et al. 2013), MONC explicitly incorporates the digital transformations introduced by social media platforms. Specifically, MONC highlights critical shifts such as the emergence of platforms as primary intermediaries, the evolution of news media entities into active opinion leaders, the increasingly empowered role of audiences in information redistribution, the direct and dynamic interactions between opinion leaders and audiences, and the critical influence of regulatory environments.

From MONC's development and systematic literature analysis, we articulate a clear, strategic research agenda for future scholarship. Essential next empirical steps include validating MONC's roles and communication patterns across diverse platforms and cultural contexts, developing precise operationalizations of complex constructs like algorithmic mediation and audience behavior, and conducting intervention-focused studies to assess platform-based strategies aimed at mitigating misinformation spread and polarization.

Broader societal and practical implications of MONC are substantial, especially for practitioners, policymakers, and platform designers. Practitioners can utilize MONC to better understand misinformation dynamics, enhancing their strategies for misinformation management. Policymakers might leverage MONC insights to develop robust regulatory frameworks that enhance transparency and accountability of digital platforms. Platform designers can use MONC to inform the development of transparent and responsible algorithms and effective user education tools, thereby fostering a healthier online information environment. Overall, MONC contributes to more effective governance and design of online platforms, directly addressing pressing societal issues surrounding misinformation, polarization, and digital literacy.

References

Abdalla Mikhaeil, C., and Baskerville, R. L. 2024. "Explaining Online Conspiracy Theory Radicalization: A Second-Order Affordance for Identity-Driven Escalation," Information Systems Journal (34:3), pp. 711-735.

Acomi, N., Siguencia, L. O., and Acomi, O. 2021. "An Appropriate Set of Skills for Limiting the Spread of Fake News," Revista Romaneasca pentru Educatie Multidimensionala (13:1), pp. 71-80.

Ahmed, S. 2021. "Who Inadvertently Shares Deepfakes? Analyzing the Role of Political Interest, Cognitive Ability, and Social Network Size," Telematics and Informatics (57).

Ahmed, W., Segui, F. L., Vidal-Alaball, J., and Katz, M. S. 2020. "Covid-19 and the "Film Your Hospital" Conspiracy Theory: Social Network Analysis of Twitter Data," Journal of Medical Internet Research (22:10), p. 8.

Ali, S., Saeed, M. H., Aldreabi, E., Blackburn, J., De Cristofaro, E., Zannettou, S., and Stringhini, G. 2021. "Understanding the Effect of Deplatforming on Social Networks," Proceedings of the 13th ACM Web Science Conference 2021, pp. 187-195.

Aljasir, S. A., and Bjanaid, W. N. 2021. "Fake News in the Era of Digital Media: What Factors Affect User Interactions with Fake News Targeting National Security?," International Journal of Future Generation Communication and Networking (14:1), pp. 21-36.

Allcott, H., and Gentzkow, M. 2017. "Social Media and Fake News in the 2016 Election," Journal of Economic Perspectives (31:2), pp. 211-235.

Allen, J., Watts, D. J., and Rand, D. G. 2024. "Quantifying the Impact of Misinformation and Vaccine-Skeptical Content on Facebook," Science (384:6699), p. eadk3451.

Anansaringkarn, P., and Neo, R. 2021. "How Can State Regulations over the Online Sphere Continue to Respect the Freedom of Expression? A Case Study of Contemporary 'Fake News' Regulations in Thailand," Information & Communications Technology Law (30:3), pp. 283-303.

Andorfer, A. 2018. "Spreading Like Wildfire: Solutions for Abating the Fake News Problem on Social Media Via Technology Controls and Government Regulation," Hastings Law Journal (69:5), pp. 1409-1431.

Anspach, N. M., and Carlson, T. N. 2022. "Not Who You Think? Exposure and Vulnerability to Misinformation," New Media & Society).

Apuke, O. D., and Omar, B. 2021. "Fake News and Covid-19: Modelling the Predictors of Fake News Sharing among Social Media Users," Telematics and Informatics (56).

Babcock, M., Beskow, D. M., and Carley, K. M. 2019. "Different Faces of False: The Spread and Curtailment of False Information in the Black Panther Twitter Discussion," Acm Journal of Data and Information Quality (11:4).

Bak-Coleman, J. B., Kennedy, I., Wack, M., Beers, A., Schafer, J. S., Spiro, E. S., Starbird, K., and West, J. D. 2022. "Combining Interventions to Reduce the Spread of Viral Misinformation," Nature Human Behaviour (6:10), pp. 1372-1380.

Bakshy, E., Messing, S., and Adamic, L. 2015. "Exposure to Ideologically Diverse News and Opinion on Facebook," Science (348:6239), pp. 1130-1132.

Balakrishnan, V., Ng, K. S., and Rahim, H. A. 2021. "To Share or Not to Share - the Underlying Motives of Sharing Fake News Amidst the Covid-19 Pandemic in Malaysia," Technology in Society (66).

Balmas, M. 2014. "When Fake News Becomes Real: Combined Exposure to Multiple News Sources and Political Attitudes of Inefficacy, Alienation, and Cynicism," Communication research (41:3), pp. 430-454.

Bambauer, D. E. 2015. "Against Jawboning," Minn. L. Rev. (100), p. 51.

Bansal, G., Weinschenk, A., and Assoc Informat, S. 2020. "Something Real About Fake News: The Role of Polarization and Social Media Mindfulness," Conference of the Association-for-Information-Systems (AMCIS), Electr Network.

Baumann, M., and Humprecht, E. 2024. "Examining the Spread of Disinformation on Facebook During the First Wave of the Covid-19 Pandemic: A Case Study in Switzerland," COMMUNICATIONS-EUROPEAN JOURNAL OF COMMUNICATION RESEARCH).

Bellutta, D., King, C., and Carley, K. M. 2021. "Deceptive Accusations and Concealed Identities as Misinformation Campaign Strategies," Computational and Mathematical Organization Theory (27:3), pp. 302-323.

Benkler, Y., Tilton, C., Etling, B., Roberts, H., Clark, J., Faris, R., Kaiser, J., and Schmitt, C. 2020. "Mail-in Voter Fraud: Anatomy of a Disinformation Campaign,").

Bennett, W. L., and Segerberg, A. 2012. "The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics," Information, communication & society (15:5), pp. 739-768.

Bergström, A., and Jervelycke Belfrage, M. 2018. "News in Social Media: Incidental Consumption and the Role of Opinion Leaders," Digital journalism (6:5), pp. 583-598.

Berry, R. A., Rosenbaum, J. E., and Corey, A. M. 2023. "Weaponising #Fakenews on Twitter: Generating Flak or Critiquing the Status Quo in the Trump Era?," Javnost-the Public), p. 17.

Bolin, G. 2012. "The Labour of Media Use: The Two Active Audiences," Information, Communication & Society (15:6), pp. 796-814.

Boulianne, S., Tenove, C., and Buffie, J. 2022. "Complicating the Resilience Model: A Four-Country Study About Misinformation," Media and Communication (10:3), pp. 169-182.

Bozdag, E., and van den Hoven, J. 2015. "Breaking the Filter Bubble: Democracy and Design," Ethics and Information Technology (17), pp. 249-265.

Bro, P., and Wallberg, F. 2014. "Digital Gatekeeping: News Media Versus Social Media," Digital journalism (2:3), pp. 446-454.

Bryanov, K., Kliegl, R., Koltsova, O., Lokot, T., Miltsov, A., Pashakhin, S., Porshnev, A., Sinyavskaya, Y., Terpilovskii, M., and Vziatysheva, V. 2023. "What Drives Perceptions of Foreign News Coverage Credibility? A Cross-National Experiment Including Kazakhstan, Russia, and Ukraine," Political Communication (40:2), pp. 115-146.

Buchanan, T., and Benson, V. 2019. "Spreading Disinformation on Facebook: Do Trust in Message Source, Risk Propensity, or Personality Affect the Organic Reach of "Fake News"?," Social Media + Society (5:4).

Buturoiu, R., Udrea, G., Oprea, D.-A., and Corbu, N. 2021. "Who Believes in Conspiracy Theories About the Covid-19 Pandemic in Romania? An Analysis of Conspiracy Theories Believers' Profiles," Societies (11:4), p. 138.

Calvillo, D. P., and Harris, J. D. 2022. "Exposure to Headlines as Questions Reduces Illusory Truth for Subsequent Headlines," Journal of Applied Research in Memory and Cognition).

Canavilhas, J., Bittencourt, M., and De Andrade, M. A. A. 2019. "Viral Content on Facebook: A Case Study on the Run-up to the Brazilian 2018 Presidential Elections," Brazilian Journalism Research (15:3), pp. 562-589.

Carrera, P. D. 2023. "Why Fake News Is Shared on Social Media an Analysis of the Peruvian Case During the Coronavirus Crisis," Question (3:74), p. 34.

Castioni, P., Andrighetto, G., Gallotti, R., Polizzi, E., and De Domenico, M. 2022. "The Voice of Few, the Opinions of Many: Evidence of Social Biases in Twitter Covid-19 Fake News Sharing," Royal Society Open Science (9:10), p. 220716.

Chan, C.-h., and Fu, K.-w. 2017. "The Relationship between Cyberbalkanization and Opinion Polarization: Time-Series Analysis on Facebook Pages and Opinion Polls During the Hong Kong Occupy Movement and the Associated Debate on Political Reform," Journal of Computer-Mediated Communication (22:5), pp. 266-283.

Chaudhuri, N., Gupta, G., and Popovic, A. 2025. "Do You Believe It? Examining User Engagement with Fake News on Social Media Platforms," TECHNOLOGICAL FORECASTING AND SOCIAL CHANGE (212).

Cheng, Y., and Luo, Y. J. 2021. "The Presumed Influence of Digital Misinformation: Examining Us Public's Support for Governmental Restrictions Versus Corrective Action in the Covid-19 Pandemic," Online Information Review (45:4), pp. 834-852.

Chibuwe, A. 2020. "Social Media and Elections in Zimbabwe: Twitter War between Pro-Zanu-Pf and Pro-Mdc-a Netizens," Communicatio-South African Journal for Communication Theory and Research (46:4), pp. 7-30.

Christner, C. 2023. "Does Credibility Become Trivial When the Message Is Right? Populist Radical-Right Attitudes, Perceived Message Credibility, and the Spread of Disinformation," Communications-European Journal of Communication Research), p. 22.

Chung, M. 2023. "Share to Stop the Harm: How Social Media Metrics Drive Sharing of Fact-Checking Messages Via First-Person Perception," Mass Communication and Society), p. 24.

Chung, M., and Kim, N. 2021. "When I Learn the News Is False: How Fact-Checking Information Stems the Spread of Fake News Via Third-Person Perception," Human Communication Research (47:1), pp. 1-24.

Chung, M., Moon, W. K., and Jones-Jang, S. M. 2023. "Ai as an Apolitical Referee: Using Alternative Sources to Decrease Partisan Biases in the Processing of Fact-Checking Messages," Digital Journalism), p. 22.

Cinelli, M., Cresci, S., Galeazzi, A., Quattrociocchi, W., and Tesconi, M. 2020a. "The Limited Reach of Fake News on Twitter During 2019 European Elections," Plos One (15:6), p. 13.

Cinelli, M., Cresci, S., Galeazzi, A., Quattrociocchi, W., and Tesconi, M. 2020b. "The Limited Reach of Fake News on Twitter During 2019 European Elections," Plos One (15:6).

Citron, D. K. 2019. The National Security Challenge of Artificial Intelligence, Manipulated Media, and 'Deep Fakes. Boston University.

Clark, B., and Stoakes, E. 2023. "Intersections of Influence Radical Conspiracist 'Alt-Media' Narratives and the Climate Crisis in Aotearoa," Pacific Journalism Review (29:1-2), pp. 12-26.

Cohen, M. S., Halewicz, V., Yildirim, E., and Kable, J. W. 2024. "Continued Influence of False Accusations in Forming Impressions of Political Candidates," PNAS NEXUS (3:11).

Colliander, J. 2019. ""This Is Fake News": Investigating the Role of Conformity to Other Users' Views When Commenting on and Spreading Disinformation in Social Media," Computers in Human Behavior (97), pp. 202-215.

Corbu, N., Bargaoanu, A., Durach, F., and Udrea, G. 2021. "Fake News Going Viral: The Mediating Effect of Negative Emotions," Media Literacy and Academic Research (4:2), pp. 58-87.

Corneille, O., Mierop, A., and Unkelbach, C. 2020. "Repetition Increases Both the Perceived Truth and Fakeness of Information: An Ecological Account," Cognition (205).

Corsi, G., Seger, E., and Héigeartaigh, S. O. 2024. "Crowdsourcing the Mitigation of Disinformation and Misinformation: The Case of Spontaneous Community-Based Moderation on Reddit," ONLINE SOCIAL NETWORKS AND MEDIA (43-44).

Coscia, M., and Rossi, L. 2020. "Distortions of Political Bias in Crowdsourced Misinformation Flagging," Journal of the Royal Society Interface (17:167).

Cousins, S. 2020. "Covid-19 Has "Devastating" Effect on Women and Girls," The Lancet (396:10247), pp. 301-302.

Das, A., and Schroeder, R. 2021. "Online Disinformation in the Run-up to the Indian 2019 Election," Information Communication & Society (24:12), pp. 1762-1778.

Daud, M., and Zulhuda, S. 2020. "Regulating the Spread of False Content Online in Malaysia: Issues, Challenges and the Way Forward," International Journal of Business and Society (21), pp. 32-48.

Daunt, K. L., Greer, D. A., Jin, H. S., and Orpen, I. 2023. "Who Believes Political Fake News? The Role of Conspiracy Mentality, Patriotism, Perceived Threat to Freedom, Media Literacy and Concern for Disinformation," Internet Research (33:5), pp. 1849-1870.

del Vas, R. S., and Navarro, J. T. 2024. "Disinformation on the Covid-19 Pandemic and the Russia-Ukraine War: Two Sides of the Same Coin?," HUMANITIES & SOCIAL SCIENCES COMMUNICATIONS (11:1).

Dennis, A. R., Fuller, R. M., and Valacich, J. S. 2008. "Media, Tasks, and Communication Processes: A Theory of Media Synchronicity," MIS quarterly), pp. 575-600.

Dennis, A. R., and Kinney, S. T. 1998. "Testing Media Richness Theory in the New Media: The Effects of Cues, Feedback, and Task Equivocality," Information Systems Research (9:3), pp. 256-274.

DiResta, R. 2024. Invisible Rulers: The People Who Turn Lies into Reality. Hachette UK.

Dobber, T., Kruikemeier, S., Votta, F., Helberger, N., and Goodman, E. P. 2023. "The Effect of Traffic Light Veracity Labels on Perceptions of Political Advertising Source and Message Credibility on Social Media," Journal of Information Technology & Politics (22:1), p. 16.

Dow, B. J., Johnson, A. L., Wang, C. S., Whitson, J., and Menon, T. 2021. "The Covid-19 Pandemic and the Search for Structure: Social Media and Conspiracy Theories," Social and Personality Psychology Compass (15:9), p. e12636.

Duggan, K. 2023. "Inside Canada's Fight with Facebook," in: Politico.

Dumitru, E. A. 2020. "Testing Children and Adolescents' Ability to Identify Fake News: A Combined Design of Quasi-Experiment and Group Discussions," Societies (10:3), p. 12.

Durani, K., Eckhardt, A., Durani, W., Kollmer, T., and Augustin, N. 2023. "Visual Audience Gatekeeping on Social Media Platforms: A Critical Investigation on Visual Information Diffusion before and During the Russo-Ukrainian War," Information Systems Journal (34:2), pp. 415-468.

Durmaz, N., and Hengirmen, E. 2022. "The Dramatic Increase in Anti-Vaccine Discourses During the Covid-19 Pandemic: A Social Network Analysis of Twitter," Human Vaccines & Immunotherapeutics (18:1), p. 2025008.

Earl, J., Maher, T. V., and Pan, J. 2022. "The Digital Repression of Social Movements, Protest, and Activism: A Synthetic Review," Science Advances (8:10), p. eabl8198.

Esteban-Navarro, M.-Á., Nogales-Bocio, A.-I., García-Madurga, M.-Á., and Morte-Nadal, T. 2021. "Spanish Fact-Checking Services: An Approach to Their Business Models," Publications (9:3), p. 38.

Etta, G., Cinelli, M., Galeazzi, A., Valensise, C. M., Quattrociocchi, W., and Conti, M. 2022. "Comparing the Impact of Social Media Regulations on News Consumption," Ieee Transactions on Computational Social Systems (10:3).

European Commission. 2023. "The Digital Markets Act: Ensuring Fair and Open Digital Markets." Retrieved 03.06.2024, from https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en

European Parliament and of the Council. 2022. "Regulation (Eu) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on Contestable and Fair Markets in the Digital Sector and Amending Directives (Eu) 2019/1937 and (Eu) 2020/1828 (Digital Markets Act) (Text with Eea Relevance)."

Farkas, J. 2023. "Fake News in Metajournalistic Discourse," Journalism Studies (24:4), pp. 423-441.

Farrington, D. P. 2007. "Advancing Knowledge About Desistance," Journal of Contemporary Criminal Justice (23:1), pp. 125-134.

Fierens, P. I., and Rêgo, L. C. 2024. "Stopping Fake News: Who Should Be Banned?," INTERNATIONAL JOURNAL OF DATA SCIENCE AND ANALYTICS (18:3), pp. 259-273.

Flintham, M., Karner, C., Bachour, K., Creswick, H., Gupta, N., and Moran, S. 2018. "Falling for Fake News: Investigating the Consumption of News Via Social Media," Proceedings of the 2018 CHI conference on human factors in computing systems, pp. 1-10.

Fourney, A., Racz, M. Z., Ranade, G., Mobius, M., Horvitz, E., and Assoc Comp, M. 2017. "Geographic and Temporal Trends in Fake News Consumption During the 2016 Us Presidential Election," ACM Conference on Information and Knowledge Management (CIKM), Singapore, SINGAPORE, pp. 2071-2074.

Gallotti, R., Valle, F., Castaldo, N., Sacco, P., and De Domenico, M. 2020. "Assessing the Risks of 'Infodemics' in Response to Covid-19 Epidemics," Nature Human Behaviour (4:12), p. 11.

George, J. F., Gupta, M., Giordano, G., Mills, A. M., Tennant, V. M., and Lewis, C. C. 2018. "The Effects of Communication Media and Culture on Deception Detection Accuracy," MIS quarterly (42:2), pp. 551-575.

Gimpel, H., Heger, S., Olenberger, C., and Utz, L. 2021. "The Effectiveness of Social Norms in Fighting Fake News on Social Media," Journal of Management Information Systems (38:1), pp. 196-221.

Goldman, E. 2021. "Content Moderation Remedies," Mich. Tech. L. Rev. (28), p. 1.

Greene, C. M., and Murphy, G. 2021. "Quantifying the Effects of Fake News on Behavior: Evidence from a Study of Covid-19 Misinformation," Journal of Experimental Psychology-Applied (27:4), pp. 773-784.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., and Lazer, D. 2019. "Fake News on Twitter During the 2016 Us Presidential Election," Science (363:6425), pp. 374-+.

Grossman, E. 2022. "Media and Policy Making in the Digital Age," Annual Review of Political Science (25), pp. 443-461.

Guess, A. M., Nyhan, B., and Reifler, J. 2020. "Exposure to Untrustworthy Websites in the 2016 Us Election," Nature Human Behaviour (4:5), pp. 472-480.

Guo, L., and Zhang, Y. Y. 2020. "Information Flow within and across Online Media Platforms: An Agenda-Setting Analysis of Rumor Diffusion on News Websites, Weibo, and Wechat in China," Journalism Studies (21:15), pp. 2176-2195.

Guo, Z., Valinejad, J., and Cho, J. H. 2022. "Effect of Disinformation Propagation on Opinion Dynamics: A Game Theoretic Approach," Ieee Transactions on Network Science and Engineering (9:5), pp. 3775-3790.

Gurba, K., Kaczmarczyk, D., and Pajchert, B. 2019. "Fake News as a Threat for News Values in Communication and Education," 13th International Technology, Education and Development Conference (INTED), Valencia, SPAIN, pp. 6937-6945.

Guzman, A. L., and Lewis, S. C. 2020. "Artificial Intelligence and Communication: A Human–Machine Communication Research Agenda," New media & society (22:1), pp. 70-86.

Hadlington, L., Harkin, L. J., Kuss, D., Newman, K., and Ryding, F. C. 2023. "Perceptions of Fake News, Misinformation, and Disinformation Amid the Covid-19 Pandemic: A Qualitative Exploration," Psychology of Popular Media (12:1), pp. 40-49.

Hamdi, S. A. 2023. "Mining Misinformation Discourse on Social Media within the 'Ideological Square'," Discourse & Society (35:3), pp. 329-344.

Harbath, K. 2023. "Yoel Roth on Hard Tradeoffs, Speaking Publicly and the Future of Trust and Safety," in: Anchor Change with Katie Harbath, K. Harbath (ed.). https://anchorchange.substack.com/.

Hartley, K., and Khuong, V. M. 2020. "Fighting Fake News in the Covid-19 Era: Policy Insights from an Equilibrium Model," Policy Sciences (53:4), pp. 735-758.

Helm, R. K., and Nasu, H. 2021a. "Regulatory Responses to 'Fake News' and Freedom of Expression: Normative and Empirical Evaluation," Human Rights Law Review (21:2), pp. 302-328.

Helm, R. K., and Nasu, H. 2021b. "Regulatory Responses to 'Fake News' and Freedom of Expression: Normative and Empirical Evaluation," Human Rights Law Review (21:2), pp. 302-328.

Himelein-Wachowiak, M., Giorgi, S., Devoto, A., Rahman, M., Ungar, L., Schwartz, H. A., Epstein, D. H., Leggio, L., and Curtis, B. 2021. "Bots and Misinformation Spread on Social Media: Implications for Covid-19," Journal of Medical Internet Research (23:5).

Himma-Kadakas, M., and Ojamets, I. 2022. "Debunking False Information: Investigating Journalists' Fact-Checking Skills," Digital Journalism (10:5), pp. 866-887.

Ho, A., and Huang, V. 2021. "Unmasking the Ethics of Public Health Messaging in a Pandemic," Journal of Bioethical Inquiry (18), pp. 549–559.

Hoffman, S. J., and Tan, C. 2013. "Following Celebrities' Medical Advice: Meta-Narrative Analysis," Bmj (347).

Hu, D. F., and Apuke, O. D. 2023. "Modeling the Factors That Stimulates the Circulation of Online Misinformation in a Contemporary Digital Age," International Journal of Human-Computer Interaction (40:15), pp. 4132-4144.

Illia, L., Zyglidopoulos, S., and Bantimaroudis, P. 2024. "Staging the Lie: The Impact of Framing and Content on the Visibility of Fake Business News," BUSINESS & SOCIETY).

Innes, M., Dobreva, D., and Innes, H. 2021. "Disinformation and Digital Influencing after Terrorism: Spoofing, Truthing and Social Proofing," Contemporary Social Science (16:2), pp. 241-255.

Iosifidis, P., and Andrews, L. 2020. "Regulating the Internet Intermediaries in a Post-Truth World: Beyond Media Policy?," International Communication Gazette (82:3), pp. 211-230.

Ippa, S., Okubo, T., and Hashimoto, M. 2024. "An Analysis of Social Bot Activity on X in Modern Japan," IEEE ACCESS (12), pp. 125800-125808.

Jang, J. W., Lee, E. J., and Shin, S. Y. 2019. "What Debunking of Misinformation Does and Doesn't," Cyberpsychology Behavior and Social Networking (22:6), pp. 423-427.

Janze, C., and Risius, M. 2017. "Automatic Detection of Fake News on Social Media Platforms," Proceedings ot the 21st Pacific Asia Conference on Information Systems:"Societal Transformation Through IS/IT", PACIS 2017: Association for Information Systems.

Jhaver, S., Birman, I., Gilbert, E., and Bruckman, A. 2019. "Human-Machine Collaboration for Content Regulation: The Case of Reddit Automoderator," ACM Transactions on Computer-Human Interaction (TOCHI) (26:5), pp. 1-35.

Jiang, B. W., and Wang, D. S. 2024. "Perception of Misinformation on Social Media among Chinese College Students," FRONTIERS IN PSYCHOLOGY (15).

Jiang, J. A., Nie, P., Brubaker, J. R., and Fiesler, C. 2023. "A Trade-Off-Centered Framework of Content Moderation," ACM Transactions on Computer-Human Interaction (30:1), pp. 1-34.

Jin, S. V., Muqaddam, A., and Ryu, E. 2019. "Instafamous and Social Media Influencer Marketing," Marketing Intelligence & Planning (37:5), pp. 567-579.

Jost, P. J., Punder, J., and Schulze-Lohoff, I. 2020. "Fake News-Does Perception Matter More Than the Truth?," Journal of Behavioral and Experimental Economics (85).

Kaiser, J., Vaccari, C., and Chadwick, A. 2022. "Partisan Blocking: Biased Responses to Shared Misinformation Contribute to Network Polarization on Social Media," Journal of Communication (72:2), pp. 214-240.

Kaplan, A. M., and Haenlein, M. 2010. "Users of the World, Unite! The Challenges and Opportunities of Social Media," Business horizons (53:1), pp. 59-68.

Karanasios, S., and Risius, M. 2024. "Algorithms Are Pushing Ai-Generated Falsehoods at an Alarming Rate. How Do We Stop This?," in: The Conversation.

Katz, E., and Lazarsfeld, P. F. 1964. Personal Influence, the Part Played by People in the Flow of Mass Communications. Transaction publishers.

Katzenbach, C. 2021. "“Ai Will Fix This”–the Technical, Discursive, and Political Turn to Ai in Governing Communication." SAGE Publications Sage UK: London, England, p. 20539517211046182.

Kaye, B., and Jackson, L. 2024. "Facebook Owner Meta Angers Australia with Plan to Stop Paying for News Content." from https://www.reuters.com/technology/meta-says-will-no-longer-pay-traditional-news-content-australia-france-germany-2024-03-01/

Kaye, D. 2019. Speech Police: The Global Struggle to Govern the Internet. Columbia Global Reports.

Kitchens, B., Johnson, S. L., and Gray, P. 2020. "Understanding Echo Chambers and Filter Bubbles: The Impact of Social Media on Diversification and Partisan Shifts in News Consumption," MIS quarterly (44:4), pp. 1619-1649.

Kluknavská, A., Eisele, O., Bartkowska, M., and Kriegler, N. 2025. "A Question of Truth: Accusations of Untruthfulness by Populist and Non-Populist Politicians on Facebook During the Covid-19 Crisis," INFORMATION COMMUNICATION & SOCIETY (28:2), pp. 258-277.

Kostakos, P., Nykanen, M., Martinviita, M., Pandya, A., and Oussalah, M. 2018. Meta-Terrorism: Identifying Linguistic Patterns in Public Discourse after an Attack.

Krasni, J. 2020. "How to Hijack a Discourse? Reflections on the Concepts of Post-Truth and Fake News," Humanities and Social Sciences Communications (7:1), pp. 1-10.

Kwanda, F. A., and Lin, T. T. C. 2020. "Fake News Practices in Indonesian Newsrooms During and after the Palu Earthquake: A Hierarchy-of-Influences Approach," Information Communication & Society (23:6), pp. 849-866.

Larson, K., and Watson, R. 2011. "The Value of Social Media: Toward Measuring Social Media Strategies," in: International Conference on Information Systems (ICIS). Shanghai, China.

Larsson, A. O. 2020. "Right-Wingers on the Rise Online: Insights from the 2018 Swedish Elections," New Media & Society (22:12), pp. 2108-2127.

Lazar, L., and Pop, M. I. 2021. "Impact of Celebrity Endorsement and Breaking News Effect on the Attention of Consumers," Studia Universitatis Vasile Goldis Arad Seria Stiinte Economice (31:3), pp. 60-74.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., and Rothschild, D. 2018. "The Science of Fake News," Science (359:6380), pp. 1094-1096.

Le, M. T. H. 2024. "The Spread of Fake News: Disclosure Willingness Role," HELIYON (10:14).

Lee, E. W. J., Bao, H. Y., Wang, Y. X., and Lim, Y. T. 2023. "From Pandemic to Plandemic: Examining the Amplification and Attenuation of Covid-19 Misinformation on Social Media," Social Science & Medicine (328), p. 15.

Lees, J., Banas, J. A., Linvill, D., Meirick, P. C., and Warren, P. 2023. "The Spot the Troll Quiz Game Increases Accuracy in Discerning between Real and Inauthentic Social Media Accounts," Pnas Nexus (2:4), p. 11.

Lepird, C. S., Ng, L. H. X., Wu, A. N., and Carley, K. M. 2024. "What News Is Shared Where and How: A Multi-Platform Analysis of News Shared During the 2022 U.S. Midterm Elections," SOCIAL MEDIA + SOCIETY (10:2).

Levens, S. M., Eltayeby, O., Aleshire, B., Nandu, S., Wesslen, R., Gallicano, T., and Shaikh, S. 2018. "Using Information Processing Strategies to Predict Contagion of Social Media Behavior: A Theoretical Model," 9th Int Conf on Appl Human Factors and Ergon (AHFE) / Int Conf on Human Factors, Software, and Syst Engn, Artificial Intelligence and Social Comp, Human Side of Serv Engn and Human Factors in Energy (HSSE), Orlando, FL, pp. 369-378.

Lewis, S. C., and Westlund, O. 2015. "Actors, Actants, Audiences, and Activities in Cross-Media News Work: A Matrix and a Research Agenda," Digital journalism (3:1), pp. 19-37.

Li, L., Aldosery, A., Vitiugin, F., Nathan, N., Novillo-Ortiz, D., Castillo, C., and Kostkova, P. 2021. "The Response of Governments and Public Health Agencies to Covid-19 Pandemics on Social Media: A Multi-Country Analysis of Twitter Discourse," Frontiers in Public Health (9).

Lima-Strong, C. 2025. "Transcript: Former Exec Sarah Wynn-Williams Testifies on Facebook's Courtship of China." Retrieved 07.05.2025, from https://www.techpolicy.press/transcript-former-exec-sarah-wynnwilliams-testifies-on-facebooks-courtship-of-china/

Linvill, D. L., Boatwright, B. C., Grant, W. J., and Warren, P. L. 2019. ""The Russians Are Hacking My Brain!" Investigating Russia's Internet Research Agency Twitter Tactics During the 2016 United States Presidential Campaign," Computers in Human Behavior (99), pp. 292-300.

Linvill, D. L., and Warren, P. L. 2020. "Troll Factories: Manufacturing Specialized Disinformation on Twitter," Political Communication (37:4), pp. 447-467.

Livingstone, S., and Helsper, E. 2010. "Balancing Opportunities and Risks in Teenagers' Use of the Internet: The Role of Online Skills and Internet Self-Efficacy," New media & society (12:2), pp. 309-329.

Llansó, E. J. 2020. "No Amount of "Ai" in Content Moderation Will Solve Filtering's Prior-Restraint Problem," Big Data & Society (7:1), p. 2053951720920686.

London Jr, J., Li, S., and Sun, H. 2022. "Seems Legit: An Investigation of the Assessing and Sharing of Unverifiable Messages on Online Social Networks," Information Systems Research (33:3), pp. 978-1001.

Lopez-Garcia, X., Vizoso, A., and Perez-Seijo, S. 2019. "Verification Initiatives in the Scenario of Misinformation. Actants for Integrated Plans with Multi-Level Strategies," Brazilian Journalism Research (15:3), pp. 614-635.

Losifidis, P., and Nicoli, N. 2020. "The Battle to End Fake News: A Qualitative Content Analysis of Facebook Announcements on How It Combats Disinformation," International Communication Gazette (82:1), pp. 60-81.

Louis, W., Thomas, E., McGarty, C., Lizzio-Wilson, M., Amiot, C., and Moghaddam, F. 2020. "The Volatility of Collective Action: Theoretical Analysis and Empirical Data," Political Psychology (41), pp. 35-74.

Lu, Y., Wu, J., Tan, Y., and Chen, J. 2022. "Microblogging Replies and Opinion Polarization: A Natural Experiment," MIS Quarterly (46:4), pp. 1901-1936.

Luo, M. F., Hancock, J. T., and Markowitz, D. M. 2022. "Credibility Perceptions and Detection Accuracy of Fake News Headlines on Social Media: Effects of Truth-Bias and Endorsement Cues," Communication Research (49:2), pp. 171-195.

Malik, A., Bashir, F., and Mahmood, K. 2023. "Antecedents and Consequences of Misinformation Sharing Behavior among Adults on Social Media During Covid-19," Sage Open (13:1), p. 14.

Maningo, C., Dacuno, J., Del Rosario, N., Dumaguit, R., Garay, C., and Villahermosa, H. 2022. "Covid-19 Social Media Infodemic: The New Media Literacy Level and the Ability to Recognize Fake News among Cebu City Young Adults," Asr Chiang Mai University Journal of Social Sciences and Humanities (9:2).

Mansoori, A., Tahat, K., Tahat, D., Habes, M., Salloum, S. A., Mesbah, H., and Elareshi, M. 2023. "Gender as a Moderating Variable in Online Misinformation Acceptance During Covid-19," Heliyon (9:9), p. 12.

Marecos, J., Shattock, E., Bartlett, O., Goiana-da-Silva, F., Maheswaran, H., Ashrafian, H., and Darzi, A. 2023. "Health Misinformation and Freedom of Expression: Considerations for Policymakers," Health Economics Policy and Law (18:2), pp. 204-217.

Matherly, T., and Greenwood, B. N. 2024. "No News Is Bad News: The Internet, Corruption, and the Decline of the Fourth Estate," Management Information Systems Quarterly (48:2), pp. 699-714.

Maweu, J. M. 2019. ""Fake Elections"? Cyber Propaganda, Disinformation and the 2017 General Elections in Kenya," African Journalism Studies (40:4), pp. 62-76.

Mazhar, B., Niu, J., Zhong, Q. Y., and Ul Haq, I. 2024. "Unraveling the Dynamics of Covid-19 Fake News Spread: The Influence of Socialization, Entertainment, and Altruism in Instant News Sharing on Social Media," TELEMATICS AND INFORMATICS REPORTS (15).

McQuail, D., and Windahl, S. 2015. Communication Models for the Study of Mass Communications. Routledge.

Méndez-Muros, S., Alonso-González, M., and Pérez-Curiel, C. 2024. "Disinformation and Fact-Checking in the Face of Natural Disasters: A Case Study on Turkey-Syria Earthquakes," SOCIETIES (14:4).

Microsoft. 2019. "Photodna." Retrieved 18.07.2023, from https://www.microsoft.com/en-us/photodna

Milan, S. 2015. "When Algorithms Shape Collective Action: Social Media and the Dynamics of Cloud Protesting," Social Media+ Society (1:2), p. 2056305115622481.

Miletskiy, V. P., Cherezov, D. N., and Strogetskaya, E. V. 2019. "Transformations of Professional Political Communications in the Digital Society (by the Example of the Fake News Communication Strategy)," IEEE Communication Strategies in Digital Society Seminar (ComSDS), Saint Petersburg Electrotechn Univ LETI, St. Petersburg, RUSSIA, pp. 121-124.

Miller, S., Menard, P., and Bourrie, D. 2024. "I'm Not Fluent: How Linguistic Fluency, New Media Literacy, and Personality Traits Influence Fake News Engagement Behavior on Social Media," Information & Management (61:2), p. 103912.

Millet, B., Tang, J. J., Seelig, M., Petit, J., and Sun, R. Y. 2024. "In Twitter We Trust(Ed): How Perceptions of Twitter's Helpfulness Influence News Post Credibility Perceptions and News Engagement," COMPUTERS IN HUMAN BEHAVIOR (155).

Morales-i-Gras, J. 2020. "Cognitive Biases in Link Sharing Behavior and How to Get Rid of Them: Evidence from the 2019 Spanish General Election Twitter Conversation," Social Media + Society (6:2).

Moravec, P., Minas, R., and Dennis, A. R. 2019. "Fake News on Social Media: People Believe What They Want to Believe When It Makes No Sense at All," MIS Quarterly (43:4), pp. 1343-1360.

Moravec, P. L., Collis, A., and Wolczynski, N. 2023. "Countering State-Controlled Media Propaganda through Labeling: Evidence from Facebook," Information Systems Research (35:3), pp. 1435-1447.

Moravec, P. L., Kim, A., Dennis, A. R., and Minas, R. K. 2022. "Do You Really Know If It's True? How Asking Users to Rate Stories Affects Belief in Fake News on Social Media," Information Systems Research (33:3), pp. 887-907.

Morejon-Llamas, N., Martin-Ramallal, P., and Micaletto-Belda, J. P. 2022. "Twitter Content Curation as an Antidote to Hybrid Warfare During Russia's Invasion of Ukraine," Profesional De La Informacion (31:3).

Mourão, R. R., and Molyneux, L. 2021. "Tweeting Outside the Lines: Normalization and Fragmentation as Political Reporters Break from the Mainstream," Journalism Practice (15:8), pp. 1089-1107.

Mourao, R. R., and Robertson, C. T. 2019. "Fake News as Discursive Integration: An Analysis of Sites That Publish False, Misleading, Hyperpartisan and Sensational Information," Journalism Studies (20:14), pp. 2077-2095.

Nekmat, E. 2020. "Nudge Effect of Fact-Check Alerts: Source Influence and Media Skepticism on Sharing of News Misinformation in Social Media," Social Media + Society (6:1).

Neyazi, T. A., Ng, A. Y. K., Kuru, O., and Muhtadi, B. 2022. "Who Gets Exposed to Political Misinformation in a Hybrid Media Environment? The Case of the 2019 Indonesian Election," Social Media + Society (8:3).

Ng, K. C., Tang, J., and Lee, D. 2021. "The Effect of Platform Intervention Policies on Fake News Dissemination and Survival: An Empirical Examination," Journal of Management Information Systems (38:4), pp. 898-930.

Nguyen, C. T. 2020. "Echo Chambers and Epistemic Bubbles," Episteme-a Journal of Individual and Social Epistemology (17:2), pp. 141-161.

Nicholas, G. 2022. "Shedding Light on Shadowbanning," in: Center for Democracy and Technology. https://cdt.org/insights/shedding-light-on-shadowbanning.

Omar, B., Apuke, O. D., and Nor, Z. M. 2023. "The Intrinsic and Extrinsic Factors Predicting Fake News Sharing among Social Media Users: The Moderating Role of Fake News Awareness," Current Psychology (43:2), pp. 1235-1247.

Osatuyi, B., and Dennis, A. R. 2024. "The Strength of Weak Ties and Fake News Believability," Decision Support Systems (184), p. 114275.

Paisana, M., Pinto-Martinho, A., and Cardoso, G. 2020. "Trust and Fake News: Exploratory Analysis of the Impact of News Literacy on the Relationship with News Content in Portugal," Communication & Society-Spain (33:2), pp. 105-117.

Pang, H., Liu, J., and Lu, J. H. 2022. "Tackling Fake News in Socially Mediated Public Spheres: A Comparison of Weibo and Wechat," Technology in Society (70).

Pantumsinchai, P. 2018. "Armchair Detectives and the Social Construction of Falsehoods: An Actor-Network Approach," Information Communication & Society (21:5), pp. 761-778.

Papanastasiou, Y. 2020. "Fake News Propagation and Detection: A Sequential Model," Management Science (66:5), pp. 1826-1846.

Parikh, S. B., Patil, V., and Atrey, P. K. 2019. "On the Origin, Proliferation and Tone of Fake News," 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR): IEEE, pp. 135-140.

Park, I., and Lee, D. 2020. "Understanding News-Sharing across Different Platforms: The Effects of Newsworthiness and Gratifications from News-Sharing," Information Research-an International Electronic Journal (25:4).

Park, M., and Chai, S. 2024. "What Makes Fake News Appeal to You? Empirical Evidence from the Tweets Related to Covid-19 Vaccines," SAGE OPEN (14:3).

Park, S., Fisher, C., Flew, T., and Dulleck, U. 2020. "Global Mistrust in News: The Impact of Social Media on Trust," Jmm-International Journal on Media Management (22:2), pp. 83-96.

Pennycook, G., Bear, A., Collins, E. T., and Rand, D. G. 2020a. "The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines without Warnings," Management Science (66:11), pp. 4944-4957.

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., and Rand, D. G. 2021. "Shifting Attention to Accuracy Can Reduce Misinformation Online," Nature (592:7855), pp. 590-+.

Pennycook, G., McPhetres, J., Zhang, Y. H., Lu, J. G., and Rand, D. G. 2020b. "Fighting Covid-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention," Psychological Science (31:7), pp. 770-780.

Pennycook, G., and Rand, D. G. 2019a. "Fighting Misinformation on Social Media Using Crowdsourced Judgments of News Source Quality," Proceedings of the National Academy of Sciences of the United States of America (116:7), pp. 2521-2526.

Pennycook, G., and Rand, D. G. 2019b. "Lazy, Not Biased: Susceptibility to Partisan Fake News Is Better Explained by Lack of Reasoning Than by Motivated Reasoning," Cognition (188), pp. 39-50.

Perel, M., and Elkin-Koren, N. 2022. Democratic Friction in Speech Governance by Ai.

Peucker, M., and Fisher, T. J. 2022. "Mainstream Media Use for Far-Right Mobilisation on the Alt-Tech Online Platform Gab," Media Culture & Society (45:2), pp. 354-372.

Pop, M. I., and Ene, I. 2019. "Influence of the Educational Level on the Spreading of Fake News Regarding the Energy Field in the Online Environment," Proceedings of the International Conference on Business Excellence (13:1), pp. 1108-1117.

Preston, S., Anderson, A., Robertson, D. J., Shephard, M. P., and Huhe, N. 2021. "Detecting Fake News on Facebook: The Role of Emotional Intelligence," Plos one (16:3), p. e0246757.

Prochaska, S., Duskin, K., Kharazian, Z., Minow, C., Blucker, S., Venuto, S., West, J. D., and Starbird, K. 2023. "Mobilizing Manufactured Reality: How Participatory Disinformation Shaped Deep Stories to Catalyze Action During the 2020 Us Presidential Election," Proceedings of the ACM on Human-Computer Interaction (7:CSCW1), pp. 1-39.

Rampersad, G., and Althiyabi, T. 2020. "Fake News: Acceptance by Demographics and Culture on Social Media," Journal of Information Technology & Politics (17:1), pp. 1-11.

Riemer, K., and Peter, S. 2021. "Algorithmic Audiencing: Why We Need to Rethink Free Speech on Social Media," Journal of Information Technology (36:4), pp. 409-426.

Risius, M., Aydinguel, O., and Haug, M. 2019. "Towards an Understanding of Conspiracy Echo Chambers on Facebook," 27th European Conference on Information Sytems (ECIS), Stockholm-Uppsala, Sweden, pp. 1-10.

Risius, M., Blasiak, K. M., Wibisono, S., and Louis, W. R. 2023. "The Digital Augmentation of Extremism: Reviewing and Guiding Online Extremism Research from a Sociotechnical Perspective," Information Systems Journal (34:3), pp. 931-963.

Rivas, J. D., and Rausseo, A. E. 2024. "Democracy and Control of Content on the Network: Special Reference to Spain," REVISTA CHILENA DE DERECHO Y TECNOLOGIA (13).

Robertson, R. E., Green, J., Ruck, D. J., Ognyanova, K., Wilson, C., and Lazer, D. 2023. "Users Choose to Engage with More Partisan News Than They Are Exposed to on Google Search," Nature (618:7964), pp. 342-348.

Ross, A. S., and Rivers, D. J. 2018. "Discursive Deflection: Accusation of "Fake News" and the Spread of Mis- and Disinformation in the Tweets of President Trump," Social Media + Society (4:2).

Ross, B., Pilz, L., Cabrera, B., Brachten, F., Neubaum, G., and Stieglitz, S. 2019. "Are Social Bots a Real Threat? An Agent-Based Model of the Spiral of Silence to Analyse the Impact of Manipulative Actors in Social Networks," European Journal of Information Systems (28:4), pp. 394-412.

Rosu, M. M., Cosmoiu, A. M., Ianole-Calin, R., and Cornoiu, S. 2023. "Stories Worth Sharing - Why Do People Spread News Online?," Online Information Review (48:3), pp. 581-600.

Rovetta, A., and Bhagavathula, A. S. 2020. "Global Infodemiology of Covid-19: Analysis of Google Web Searches and Instagram Hashtags," Journal of Medical Internet Research (22:8), p. 9.

Salge, C. A. d. L., Karahanna, E., and Thatcher, J. B. 2022. "Algorithmic Processes of Social Alertness and Social Transmission: How Bots Disseminate Information on Twitter," Mis Quarterly (46:1).

San Martin, J., Drubi, F., and Perez, D. R. 2020. "Uncritical Polarized Groups: The Impact of Spreading Fake News as Fact in Social Networks," Mathematics and Computers in Simulation (178), pp. 192-206.

Santosa, H. P., Hasfi, N., and Lukmantoro, T. 2018. "Digital Media Unequality During the 2014th Indonesian Presidential Election," 3rd International Conference on Energy, Environmental and Information System (ICENIS) - Strengthening Planning and Implementation Energy, Environment, and Information System Toward Low Carbon Society, Semarang, INDONESIA.

Sauer, M. A., Truelove, S., Gerste, A. K., and Limaye, R. J. 2021. "A Failure to Communicate? How Public Messaging Has Strained the Covid-19 Response in the United States," Health security (19:1), pp. 65-74.

Schinello, S. 2024. "New (Digital) Media in Creative Society: Ethical Issues of Content Moderation," FILOSOFIJA-SOCIOLOGIJA (35:1), pp. 67-74.

Schlag, G. 2023. "European Union's Regulating of Social Media: A Discourse Analysis of the Digital Services Act," Politics and Governance (11:3), pp. 168-177.

Schneiders, P. 2023. "News from the User's Perspective: With Naivety to Validity," Digital Journalism), p. 22.

Shaban, S. 2020. "Teenagers, Terrorism, and Technopanic: How British Newspapers Framed Female Isis Recruits as Victims of Social Media," International Journal of Communication (14), p. 21.

Shannon, C. E. 2001. "A Mathematical Theory of Communication," ACM SIGMOBILE mobile computing and communications review (5:1), pp. 3-55.

Sheen, G. C. H. 2024. "Punishment Trumps Warning in Prebunking Misinformation: Evidence from Survey Experiments in Taiwan," JOURNAL OF ASIAN AND AFRICAN STUDIES).

Shin, J., Jian, L., Driscoll, K., and Bar, F. 2018. "The Diffusion of Misinformation on Social Media: Temporal Pattern, Message, and Source," Computers in Human Behavior (83), pp. 278-287.

Shu, K., Wang, S. H., Liu, H., and Ieee. 2018. Understanding User Profiles on Social Media for Fake News Detection.

Sievi, L., and Pawelec, M. 2025. "(How) Should Security Authorities Counter False Information on Social Media in Crises? A Democracy-Theoretical and Ethical Reflection," INTERNATIONAL JOURNAL OF DISASTER RISK REDUCTION (116).

Slutskiy, P., and Boonchutima, S. 2022. "Credibility of the Official Covid Communication in Thailand: When People Stop Believing the Government," American Behavioral Scientist).

Sommariva, S., Vamos, C., Mantzarlis, A., Dao, L. U. L., and Tyson, D. M. 2018. "Spreading the (Fake) News: Exploring Health Messages on Social Media and the Implications for Health Professionals Using a Case Study," American Journal of Health Education (49:4), pp. 246-255.

Srivastava, A., Kannan, R., Chelmis, C., and Prasanna, V. K. 2019a. "Recant: Network-Based Recruitment for Active Fake News Correction," IEEE International Conference on Big Data (Big Data), Los Angeles, CA, pp. 940-949.

Srivastava, A., Kannan, R., Chelmis, C., and Prasanna, V. K. 2019b. "Recant: Network-Based Recruitment for Active Fake News Correction," in 2019 Ieee International Conference on Big Data, C. Baru, J. Huan, L. Khan, X.H. Hu, R. Ak, Y. Tian, R. Barga, C. Zaniolo, K. Lee and Y.F. Ye (eds.). pp. 940-949.

Srivastava, K. C., Shrivastava, D., Chhabra, K. G., Naqvi, W., and Sahu, A. 2020. "Facade of Media and Social Media During Covid-19: A Review," Int. J. Res. Pharm. Sci. (11:SPL1), pp. 142-149.

Starbird, K. 2019. "Disinformation's Spread: Bots, Trolls and All of Us," Nature (571:7766), p. 449.

Starbird, K., DiResta, R., and DeButts, M. 2023. "Influence and Improvisation: Participatory Disinformation During the 2020 Us Election," Social Media+ Society (9:2), p. 20563051231177943.

Stenberg, M. 2024. "Google Is Paying Publishers to Test an Unreleased Gen Ai Platform," in: Adweek.

Sterrett, D., Malato, D., Benz, J., Kantor, L., Tompson, T., Rosenstiel, T., Sonderman, J., and Loker, K. 2019. "Who Shared It?: Deciding What News to Trust on Social Media," Digital Journalism (7:6), pp. 783-801.

Strand, C., and Svensson, J. 2019. "?Fake News? On Sexual Minorities Is ?Old News?: A Study of Digital Platforms as Spaces for Challenging Inaccurate Reporting on Ugandan Sexual Minorities," African Journalism Studies (40:4), pp. 77-95.

Sun, R. Y., Li, C., Millet, B., Ali, K. I., and Petit, J. 2022. "Sharing News with Online Friends: A Study of Network Homophily, Network Size, and News Type," Telematics and Informatics (67).

Suttle, R., Hogan, S., Aumaugher, R., Spradling, M., Merrigan, Z., and Straub, J. 2021. "University Community Members' Perceptions of Labels for Online Media," Future Internet (13:11), p. 281.

Tan, W. K., and Hsu, C. Y. 2023. "The Application of Emotions, Sharing Motivations, and Psychological Distance in Examining the Intention to Share Covid-19-Related Fake News," Online Information Review (47:1), pp. 59-80.

Tandoc, E. C., and Kim, H. K. 2023. "Avoiding Real News, Believing in Fake News? Investigating Pathways from Information Overload to Misbelief," Journalism (24:6), pp. 1174-1192.

Tandoc, E. C., Lee, J., Chew, M., Tan, F. X., and Goh, Z. H. 2021. "Falling for Fake News: The Role of Political Bias and Cognitive Ability," Asian Journal of Communication (31:4), pp. 237-253.

Tandoc Jr, E. C., and Vos, T. P. 2016. "The Journalist Is Marketing the News: Social Media in the Gatekeeping Process," Journalism practice (10:8), pp. 950-966.

Tantau, A., Pelau, C., and Pop, M. 2018. "Fake News in the Energy Industry," BASIQ International Conference on New Trends in Sustainable Business and Consumption, Heidelberg, GERMANY, pp. 204-210.

Tian, Y., and Willnat, L. 2025. "From News Disengagement to Fake News Engagement: Examining the Role of News-Finds-Me Perceptions in Vulnerability to Fake News through Third-Person Perception," COMPUTERS IN HUMAN BEHAVIOR (162).

Till, C. 2021. "Propaganda through 'Reflexive Control'and the Mediated Construction of Reality," New Media & Society (23:6), pp. 1362-1378.

Tolentino, D., and Sung, M. 2023. "How the East Palestine Train Derailment Fueled Fear on Tiktok," in: NBC News.

Torreblanca, J. I. 2023. "Social Networks and Democracy: Problems and Dilemmas of Regulating the Digital Ecosystem," Siyasal-Journal of Political Sciences (32:1), pp. 15-33.

Treen, K. M. d. I., Williams, H. T., and O'Neill, S. J. 2020. "Online Misinformation About Climate Change," Wiley Interdisciplinary Reviews: Climate Change (11:5), p. e665.

Tufekci, Z. 2014. "Engineering the Public: Big Data, Surveillance and Computational Politics," First Monday (19:7).

Tully, M., and Singer, J. B. 2023. "Fact-Checking the Covid-19 Infodemic in Sub-Saharan Africa," AFRICAN JOURNALISM STUDIES (44:2), pp. 97-115.

Turel, O. 2023. "Polarized Sharing of Fake News on Social Media: The Complex Roles of Partisan Identification and Gender," Behaviour & Information Technology (43:11), pp. 2424-2441.

Umejei, E., Ajaero, I., Samuels, G. E., Oloruntola, S., and Odebiyi, S. D. 2024. "Fake News in Nigeria: Global Phenomenon, Local Realities," CURRENT PSYCHOLOGY (43:29), pp. 24482-24496.

Vaast, E., Safadi, H., Lapointe, L., and Negoita, B. 2017. "Social Media Affordances for Connective Action: An Examination of Microblogging Use During the Gulf of Mexico Oil Spill," MIS quarterly (41:4), pp. 1179-1206.

Valdez, A. C., and Ziefle, M. 2018. "Understanding Users Perceptions of Fake News and the Effectiveness of Fact-Checking Badges," 20th Congress of the International-Ergonomics-Association (IEA), Florence, ITALY, pp. 469-477.

Valenzuela, S., Muniz, C., and Santos, M. 2022. "Social Media and Belief in Misinformation in Mexico: A Case of Maximal Panic, Minimal Effects?," International Journal of Press-Politics).

Van der Linden, S., Leiserowitz, A., Rosenthal, S., and Maibach, E. 2017. "Inoculating the Public against Misinformation About Climate Change," Global challenges (1:2), p. 1600008.

van Der Linden, S., Roozenbeek, J., and Compton, J. 2020. "Inoculating against Fake News About Covid-19," Frontiers in Psychology (11), p. 7.

Van Dijck, J. 2009. "Users Like You? Theorizing Agency in User-Generated Content," Media, culture & society (31:1), pp. 41-58.

Van Heekeren, M. 2020. "The Curative Effect of Social Media on Fake News: A Historical Re-Evaluation," Journalism Studies (21:3), pp. 306-318.

Verma, A., and Nayak, J. K. 2024. "The Power of Perception: How Persuasion Knowledge and Perceived Deception in Advertisement Impact Brand-Related Fake News Adoption," GLOBAL KNOWLEDGE MEMORY AND COMMUNICATION).

Vu, H. T., and Saldaña, M. 2021. "Chillin'effects of Fake News: Changes in Practices Related to Accountability and Transparency in American Newsrooms under the Influence of Misinformation and Accusations against the News Media," Journalism & mass communication quarterly (98:3), pp. 769-789.

Waruwu, B. K., Tandoc, E. C., Duffy, A., Kim, N., and Ling, R. 2021. "Telling Lies Together? Sharing News as a Form of Social Authentication," New Media & Society (23:9), pp. 2516-2533.

Weber, D., Falzon, L., Mitchell, L., and Nasim, M. 2022. "Promoting and Countering Misinformation During Australia's 2019–2020 Bushfires: A Case Study of Polarisation," Social Network Analysis and Mining (12:1), p. 64.

Wenzel, M., Stasiuk-Krajewska, K., Macková, V., and Turkova, K. 2023. "The Penetration of Russian Disinformation Related to the War in Ukraine: Evidence from Poland, the Czech Republic and Slovakia," International Political Science Review (45:2), pp. 192-208.

Westley, B. H., and MacLean Jr, M. S. 1957. "A Conceptual Model for Communications Research," Journalism Quarterly (34:1), pp. 31-38.

White, D. M. 1950. "The "Gate Keeper": A Case Study in the Selection of News," Journalism quarterly (27:4), pp. 383-390.

Wilson, T., and Starbird, K. 2021. "Cross-Platform Information Operations: Mobilizing Narratives & Building Resilience through Both'big'&'Alt'tech," Proceedings of the ACM on Human-Computer Interaction (5:CSCW2), pp. 1-32.

Wired. 2018. "Youtube Keeps Deleting Evidence of Syrian Chemical Weapon Attacks." Retrieved 08.05.2025, from https://www.wired.com/story/chemical-weapons-in-syria-youtube-algorithm-delete-video/

Wolfswinkel, J. F., Furtmueller, E., and Wilderom, C. P. 2013. "Using Grounded Theory as a Method for Rigorously Reviewing Literature," European journal of information systems (22:1), pp. 45-55.

Wong, R. Y. M., Cheung, C. M., Xiao, B., and Thatcher, J. B. 2021. "Standing up or Standing By: Understanding Bystanders' Proactive Reporting Responses to Social Media Harassment," Information Systems Research (32:2), pp. 561-581.

Wood, A. K., and Ravel, A. M. 2018. "Fool Me Once: Regulating "Fake News" and Other Online Advertising," Southern California Law Review (91:6), pp. 1223-1278.

Wu, Y. Y., Kuru, O., Campbell, S. W., and Baruh, L. 2023. "Explaining Health Misinformation Belief through News, Social, and Alternative Health Media Use: The Moderating Roles of Need for Cognition and Faith in Intuition," Health Communication (38:7), pp. 1416-1429.

Xiao, X. Z. 2022. "Not Doomed: Examining the Path from Misinformation Exposure to Verification and Correction in the Context of Covid-19 Pandemic," Telematics and Informatics (74).

Xu, W. W., Sang, Y., and Kim, C. 2020. "What Drives Hyper-Partisan News Sharing: Exploring the Role of Source, Style, and Content," Digital Journalism (8:4), pp. 486-505.

Yan, G., Pegoraro, A., and Watanabe, N. M. 2021. "Examining Ira Bots in the Nfl Anthem Protest: Political Agendas and Practices of Digital Gatekeeping," Communication & Sport (9:1), pp. 88-109.

Yesilada, M., and Lewandowsky, S. 2022. "Systematic Review: Youtube Recommendations and Problematic Content," Internet policy review (11:1), p. 1652.

Yoshikawa, K., Awa, T., Kusano, R., Ichino, M., Yoshiura, H., and Soc, I. C. 2021. "Reliability-Disguised Attacks on Social Network to Accelerate Fake News Dissemination," 21st IEEE International Conference on Software Quality, Reliability and Security (QRS), Hainan, PEOPLES R CHINA, pp. 841-848.

Youtube. 2019. "How Content Id Works.", 18.07.2023, from https://support.google.com/youtube/answer/2797370

Yu, W. T., and Yan, Q. 2024. "Reasons to Fight: Preliminary Results on Motivations to Combat Fake News," INFORMATION RESEARCH-AN INTERNATIONAL ELECTRONIC JOURNAL (29:2), pp. 30-49.

Zannettou, S., Caulfield, T., De Cristofaro, E., Kourtellis, N., Leontiadis, I., Sirivianos, M., Stringhini, G., Blackburn, J., and Acm. 2017. "The Web Centipede: Understanding How Web Communities Influence Each Other through the Lens of Mainstream and Alternative News Sources," ACM Internet Measurement Conference (IMC), London, ENGLAND, pp. 405-418.

Zecchinon, P., and Standaert, O. 2025. "The War in Ukraine through the Prism of Visual Disinformation and the Limits of Specialized Fact-Checking. A Case-Study at Le Monde," DIGITAL JOURNALISM (13:1), pp. 61-79.

Zhang, C., Cao, T., and Ali, A. 2022a. "Investigating the Role of Perceived Information Overload on Covid-19 Fear: A Moderation Role of Fake News Related to Covid-19," Frontiers in Psychology (13).

Zhang, Z., Akhter, S., Al-Abyadh, M. A., and Cong, P. T. 2022b. "Determinants of Unverified News Sharing on Social Media and Its Effects on Corporate Image," Frontiers in Psychology (13).

Zhang, Z. Q., Cheng, Z. C., Gu, T. F., and Zhang, Y. X. 2024. "Determinants of Users' Unverified Information Sharing on Social Media Platforms: A Herding Behavior Perspective," ACTA PSYCHOLOGICA (248).

Zhou, H., Lu, Y. B., Zhao, L., Wang, B., and Li, T. 2023. "Effective Reporting System to Encourage Users' Reporting Behavior in Social Media Platforms: An Empirical Study Based on Structural Empowerment Theory," Behaviour & Information Technology (43:14), pp. 3490-3509.

Zimmer, F., and Reich, A. 2018. "What Is Truth? Fake News and Their Uncovering by the Audience," in Proceedings of the 5th European Conference on Social Media, V. Cunnane and N. Corcoran (eds.). pp. 374-381.

Zimmermann, D., Noll, C., Grasser, L., Hugger, K. U., Braun, L. M., Nowak, T., and Kaspar, K. 2022. "Influencers on Youtube: A Quantitative Study on Young People's Use and Perception of Videos About Political and Societal Topics," Current Psychology (41:10), pp. 6808-6824.

Zimmermann, F., and Kohring, M. 2020. "Mistrust, Disinforming News, and Vote Choice: A Panel Survey on the Origins and Consequences of Believing Disinformation in the 2017 German Parliamentary Election," Political Communication (37:2), pp. 215-237.

Zuckerberg, M. 2018. "A Blueprint for Content Governance and Enforcement." Facebook Notes Retrieved July 12, 2021, 2021, from https://www.facebook.com/notes/751449002072082/

Chapter I
Changes of Communication

Paper II
Endorse the Source –
The Impact of Information Assessment on News Sharing Behavior

Maximilian Haug
Neu-Ulm University of Applied Sciences

Proceedings of the *Americas Conference on Information Systems (AMCIS), 2019*,
Cancún, Mexico
https://aisel.aisnet.org/amcis2019/virtual_communities/virtual_communities/17

Chapter II
Interaction

Paper III

The New Media(tion):
The Impact of Social Media Communication Characteristics on the Believability of News

Marten Risius
Neu-Ulm University of Applied Sciences

Maximilian Haug
Neu-Ulm University of Applied Sciences

Heiko Gewald
Neu-Ulm University of Applied Sciences

Jason Bennett Thatcher
University of Colorado Boulder

Unpublished

The New Media(tion):
The Impact of Social Media Communication Characteristics on the Believability of News

Abstract

Social media is replacing the institutional press as gatekeeper who provides access to audiences and builds narratives. This creates an information void that is seized by alternative media, foreign actors, or nefarious politicians. This article describes three ways in which social media has changed the information ecosystem for news media: broad engagement opportunities for all actors, the redistribution of news stories, and the immediately discernable distinction between messengers and sources of information. We then explore how these social media communication characteristics affect the believability of news stories online. To that end, we conduct two online experiments (N_1 = 120, N_2 = 900) where we simulate immediate (study 1) and mediated interactions (study 2) with partisan news media, politician, and user accounts. In the immediate scenario, participants were exposed to news stories directly shared by these accounts. In the mediated study, one of these accounts (messenger) shared news stories from another account (source). While the results show that the perceived similarity with an actor positively impacts the believability of the news they share, we also find that this relationship is not uniform and instead depends on various moderating circumstances: (1) news media is not as affected by the perceived similarity as politicians, (2) "mixed-actor-amplification-effect" as the increase in news believability when a similar messenger shares news from a dissimilar source, and (3) the believability-boosting impact of concatenating pairs of similar sources and messengers. We discuss the theoretical and practical implications of these findings for news media to fill the information void and positively impact the public.

Keywords

news media, social media communication, similarity-attraction paradigm, confirmation bias, mixed-actor-amplification-effect, online experiment, news believability

Introduction

Recent events such as the US presidential election show that the traditional media is no longer the main source of news for the public. For example, Joe Rogan's podcast interviewing Donald Trump reached 47 million viewers (PowerfulJRE 2024), outperforming prime-time cable news shows, such as CNN, whose shows average of 830,000 total viewers (Mwachiro 2024). Similarly, Kamala Harris connected with

more young women through her appearance on the podcast "Call Her Daddy" (Egan 2024) than through shows like 60 Minutes and The View combined (Khalid et al. 2024). The growth of podcasts has been accompanied by a decline in trust in the institutional press (Luo et al. 2022; Valdez and Ziefle 2018) and the decline of readership of local news media (Heese et al. 2022). The implications of the decline of traditional media have been severe, ranging from reduced advertising revenues (Gao et al. 2020), to the proliferation of disinformation (Kim and Dennis 2019; Kitchens et al. 2020), and shocking, such as increasingly unchecked corruption of public officials (Matherly and Greenwood 2024). The void created by the decline of traditional media poses a national security risk, as it has opened the door for hostile governments to exploit information gaps by paying off social media influencers to spread disinformation (Ryan et al. 2022).

In place of traditional media, a social-media driven network of content delivery has reshaped how news is shared with readers. Historically, institutional media (e.g., newspapers, television, and radio) served as gatekeepers, connecting advocates (e.g., politicians) with audiences while enforcing journalistic quality standards (e.g., verifying information from multiple independent sources) and tailoring news to the interests of the audience (Westley and MacLean Jr 1957). Currently, social media, (e.g., apps, websites, and online platforms), that do not enforce journalistic standards, filter, curate, and disseminate news to wide audiences (Riemer and Peter 2021). This dis- and reintermediation (henceforth summatively referred to as disintermediation) of communication has unleashed the rapid exchange of information and enabled may-to-many communication between politicians, users, and media (Larson and Watson 2011). Social media platforms with their algorithms and content management systems (Lewis and Westlund 2015), largely determine the information consumption of their users (Kitchens et al. 2020). As a result, even institutional media view social media as gatekeepers to disseminate verified news, shifting the distribution of power among actors participating in the online information ecosystem (Grinberg et al. 2019b).

Research has begun to investigate how disintermediated online information ecosystems affect the institutional media's role ability to engage with news audiences. For example, research has found negative side-effects of journalists who engage directly with audiences online (Lee 2015), negative reputational impact of misinformation warnings on perceptions of authentic news (van der Meer et al. 2023), or a demonstrable push toward online niche news sources that are divorced from journalistic standards and norms (Reese 2020). As news articles are increasingly distributed on social media platforms, the credibility of the institutional media news source that originally created news have been diluted (Sterrett et al. 2019). Consequently, we know that social media has compromised institutional media's ability to shape the public discourse. Instead, of trained journalists, in many cases, new actors with many attributes (e.g., anxiety, illiteracy, source ambiguity) and motivations (e.g., prior beliefs) create and drive the dissemination of news, which can translate into harmful consequences (e.g., profit erosion, reputational damage, loss of human life) (Muhammed and Mathew 2022).

Despite changes evoked by disintermediation of online information ecosystems, very little work has devoted attention to the downstream effects of disintermediation and introduction of new actors on networked communication. Rarely do studies consider the dynamics of online information ecosystems

in which news is disseminated (Sterrett et al. 2019). Further, existing social media research typically does not distinguish between the actor who originally created a post and the messenger who shares it (Kim and Dennis 2019). To address this gap, this study directs attention to the different types of actors (i.e., users, news media, politicians), the emerging communication role they play in disintermediated communication (i.e., source, messenger), and the form in which they share information (i.e., immediate, mediated) on the believability of news in disintermediated social media network communication. Thereby, we investigate how characteristics of social media communication influence the believability of news?

To address this research question, we explain how social media transformed the information ecosystem, altering the gatekeeper function of the institutional press, and its consequences for the believability of news media. We conceptualize the main changes in reference to the traditional model of mass communication (Westley and MacLean Jr 1957) in terms of (1) liberated communication for all actors (i.e., news media, politicians, users), (2) the opportunity for the redistribution of news through mediated communication, and (3) the resulting differentiation between messengers and sources. In two online experiments with 1,232 participants, we examine the effects of these changes on the believability of news stories as a function of the perceived similarity of the actors sharing the news stories.

Our study yields three main findings: First, the confirmation bias effect of perceived similarity is not uniform, but is determined by the actor sharing the news – i.e., especially for politicians but not for news media. Second, we find a "mixed-actor-amplification-effect" whereby a similar messenger sharing news from a seemingly dissimilar source appears to be the most believable. Third, the concatenation of similar actors in the communication chain increases the believability of news. Finally, we discuss how these changes in the role of institutional media and the online information ecosystem affect society at the intersection between business, technology, and public policy.

In the remainder of this paper, we discuss how social media platforms have changed the role of institutional media in the information ecosystem. We explain that certain communication characteristics of social media (i.e., the predetermined actor types, the mediated communication process, and the resulting emergent roles of messenger and source) warrant further consideration with respect to the believability of news. Then, we describe our research model, our experimental design, its execution and the empirical analysis. Finally, we discuss the results and conclude with implications.

Theoretical Background: The New Media(tion)

We begin by illustrating why it is crucial to consider how social media communication changed the way information is disseminated in society. We argue social media platforms provide a distinctly different ecosystem for news propagation when compared to traditional forms of media. To explain how social media characteristics transform the ecosystem for the dissemination of news, we conceptualize how social media communication characteristics differ from traditional forms of communication.

Historically, the communication process has been much simpler and more controlled. Communication was conceptualized by (Westley and MacLean Jr 1957). In their seminal model of communication, which disentangles forms of interpersonal and mass communication to describe how information is filtered before reaching a target audience. In particular, they distinguish between advocates (e.g., politicians), as any form of purposive promoter, who must rely on gatekeepers (e.g., news organizations) in order to reach audiences through their channel of mass communication (e.g., newspaper, television) (McQuail and Windahl 2015). As brokers, gatekeepers traditionally filter information in two ways by selecting advocates they deem to be of interest to their audience and by determining the quality and frame of the information they send to the audience (i.e., one-to-many, mass communication). The audience then shares and discusses this filtered information with others in their immediate environment (i.e., one-to-one, interpersonal communication). Thereby, gatekeepers serve as information brokers who enforce journalistic standards to ensure the quality of news, while the audiences' confirmation biases guide their selection of gatekeepers and their interpersonal interactions (Figure 1).

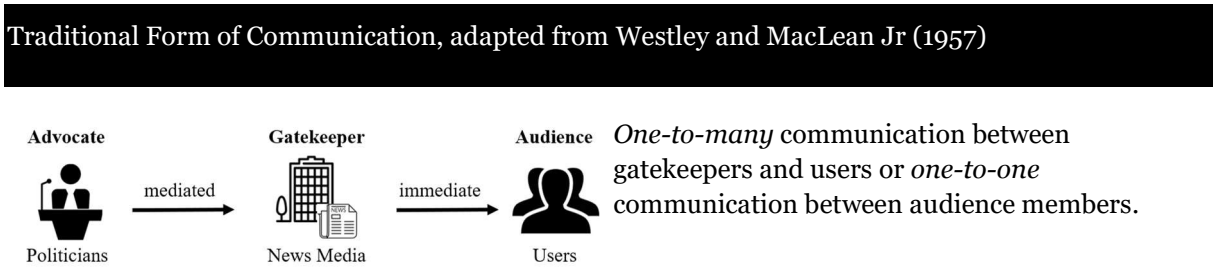


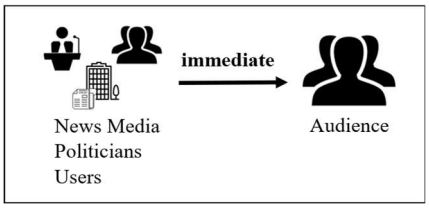
Figure 1. Traditional pattern of communication between actors.

People increasingly consume news through social media rather than through websites controlled by traditional media (Aubin and Liedke 2024; West 2017). In doing so, social media platforms increase the complexity of communication by extending traditional one-to-one and one-to-many exchanges to many-to-many communication (Larson and Watson 2011; McFarland and Ployhart 2015; Treem and Leonardi 2012). Any person or organization can freely connect with any number of other users around the world. For example, a right-wing politician from Wisconsin can freely disseminate a fabricated report about the high crime rates of refugees in Germany to a radical audience of followers in the southern United States. This information can then be picked up and spread to a US- or world-wide audience of friends and followers. All of this without the need for verification or the need to convince media organizations to pick up and disseminate the message. This information void is seized, for example, by fake or hyperpartisan alternative media who polarize the societal discourse (Linvill and Warren 2020; Vargo et al. 2018), by foreign actors who create a propagandistic narrative to corrupt the information ecosystem (Ryan et al. 2022), or by politicians for nefarious extremist agenda-setting (Starbird et al. 2023). Related research describes the impact of the continuously diminishing gatekeeping role of the institutional press, for example, on an increase in corruption (Matherly and Greenwood 2024), economic losses (Muhammed and Mathew 2022), or the public erosion of trust in information (van der Meer et al. 2023).

In this study, we aim to build an understanding of the Internet-driven disintermediation of the information ecosystem, the role of institutional media, and the promotion of news among selected audiences (Kitchens et al. 2020). Specifically, it introduced three changes that challenged the primacy of the institutional media in informing the public.

First, social media disrupts the gatekeeping function of institutional media by expanding the ability of advocates and users to immediately engage with a broad audience (Figure 2). Social media communication allows people to bypass traditional gatekeepers with their information filters and immediately engage with other users (Leong et al. 2019). By becoming part of online user networks with access to bidirectional exchange, advocates multiply their ability to disseminate information to a susceptible audience with similar interests (Sahelices-Pinto and Rodríguez-Santos 2014). The immediate form of interaction liberates advocates from the traditional filtering criteria of the gatekeeping institutional media (e.g., fact-checking, editorial judgment) and allows for direct communication and spread of news (Allcott and Gentzkow 2017; Westley and MacLean Jr 1957). Also, it allows users to engage in a virtually unlimited number of instant exchanges with other users (McFarland and Ployhart 2015) as well as redistribute information in enclaves of people who often share congruent opinions (Nguyen 2020; Spohr 2017). Thereby, social media enhances advocates' and users' ability to compete with institutional media for immediate information exchanges with audiences. Under conditions where people can engage with anyone without being constrained by content quality filters, the role of institutional media as a watchdog for society at large diminishes while conspiracy theories and false news campaigns can thrive (Linvill and Warren 2020).

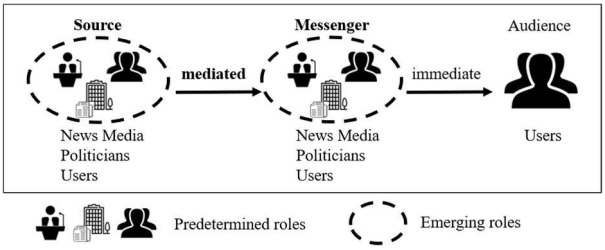
Expanded Opportunities for Immediate Exchanges on Social Media



Additional immediate *one-to-many* communication opportunities for advocates and users compete with institutional media's traditional primacy to inform the public.

Figure 2. Users and politicians competing with institutional media for exchanges with the audience on social media.

Second, social media transforms the information ecosystem by creating means for the broad redistribution of messages (e.g., sharing, retweeting) (boyd et al. 2010). The combination of mediated and immediate exchange creates a networked many-to-many communication process (Figure 3). By allowing anyone to pick up and redistribute messages among a broad audience of followers, social media allows users and advocates to also function as gatekeepers in addition to institutional media (Pal et al. 2019; Shin et al. 2018; Sommariva et al. 2018). At the same time, this creates opportunities for hyperpartisan new media (Vargo et al. 2018), fringe political advocates (Wang et al. 2020), or botnets (Salge et al. 2022) to induce political discord and gatekeep information despite feeling no obligation to fact check or vouch for the veracity of news (Bradshaw et al. 2020; Golovchenko et al. 2020). This unfettered creation and redistribution of news can be particularly powerful, given institutional media's comparatively slow and arduous quality assurance processes of institutional media (Shao et al. 2018b).



The redistribution of information in *mediated* many-to-many communication introduces the distinction between the emerging roles of *sources* and *messengers*.

Figure 3. Mediated and immediate exchanges with various gatekeepers together with the emerging roles of sources and messengers in networked communication.

Third, another implication of networked communication is that it separates predetermined (e.g., news media, politician, user) roles from emerging roles (e.g., messenger, source) according to their behavior, which IS scholars call connective affordances (Vaast et al. 2017). Emerging roles are fluid and determined by the patterns of feature use (e.g., posting, sharing). On social media, we can distinguish between source and messenger in the mediated communication condition. A source is the original creator of a message while the messenger is the person who immediately engages with the recipient (Ruggiero 2015). As an artefact of the traditional communication process, research on social media commonly confounds these roles (Janze and Risius 2017; Kim and Dennis 2019; Kim et al. 2019) or compares the messenger against the platform-external information origin as a source (George et al. 2014; Oeldorf-Hirsch and DeVoss 2020). These studies find that the perceived believability of a news article is more dependent on the trust in the origin of the information than in the messenger (George et al. 2014; Oeldorf-Hirsch and DeVoss 2020), but the messenger can dilute the impact of the source credibility (Sterrett et al. 2019). Given the ever-diminishing role of institutional media as a gatekeeper, it becomes imperative to better understand the different roles of messengers and sources in an online ecosystem where the news media have lost their grip on the distribution of information.

Research Model

We base our investigation of the transformative effects of social media on the dissemination and consumption of news on the similarity-attraction paradigm. The similarity-attraction paradigm explains that the perceived similarity to an actor shapes the believability of news. According to the paradigm, people experience a confirmation bias when they are persuaded by an actor whom they perceive as similar to themselves compared to an actor they perceive as dissimilar (Ruijten 2020; Tajfel 1982; Wilson and Sherrell 1993). Perceived similarity to actors can make people more likely to seek them out, engage with, and share information from them more readily (Lord et al. 1979). Perceived similarity does this because it reaffirms and strengthens people's confidence in previously held beliefs and makes it more difficult for them to accept alternative views (Mutahi and Kimari 2020). According to the similarity-attraction paradigm, knowledge of political affiliation would be expected to evoke feelings of

similarity and shape the perceptions of the information that they share (Roth et al. 2017). With respect to social media, there is evidence that disclosing political party affiliation does indeed evoke feelings of similarity (Pennycook and Rand 2019), which, for example, has been found to bias interviewing decisions in favor of applicants who are more similar to those empowered to make hiring decisions (Roth et al. 2019; Wade et al. 2020). The research model (Figure 4) of this study builds on the similarity-attraction paradigm and integrates multi-theoretic arguments that explain the effects of social media on the online information ecosystem, the role of institutional media, and the consumption of news.

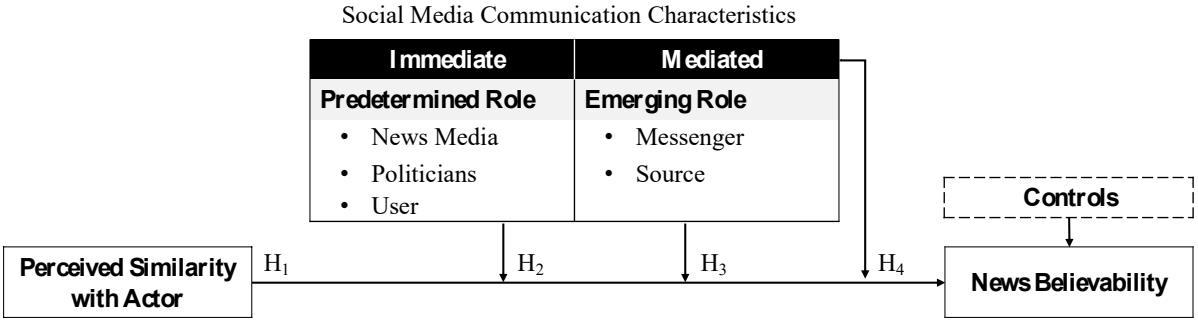


Figure 4. Research model predicting the believability of news stories by integrating the characteristics of social media communication with the similarity-attraction paradigm.

Believability of False News when Shared by Similar Actors

Social media encourages connections with users that are similar to oneself (Bakshy et al. 2015) leading, for example, to enclaves of like-minded people who share congruent views while actively excluding or discrediting dissimilar views (Nguyen 2020; Spohr 2017). For example, an analysis of 238,943 personal user networks shows great network homophily and social polarization on Twitter based on the similarity of political interests (Boutyline and Willer 2017). Similarly, content similarity algorithms are particularly good at identifying new interesting people with shared interests, as opposed to relationship-based algorithms that are better at identifying offline friends (Chen et al. 2009). Within online enclaves, social media users' opinions on a particular issue are most influenced by commenters with whom they identify with (Neubaum et al. 2018; Walther et al. 2010). Given that social media allows people to feel similar even to complete strangers (Heavey et al. 2020; Tsai and Men 2017), social media platforms become powerful tools for inducing confirmation biases by exploiting real or perceived similarity with other actors.

These confirmation bias effects of perceived similarity of actors translate into how users consume news on social media. Anspach (2017), for example, finds that social media users are more likely to engage with certain news stories when their friends and family post them, while suggestions from perceived strangers have no effect. Greater perceived similarity to another user also reduces the probability of flagging false articles that they post (Coscia and Rossi 2020). While some research has examined the confirmation bias effects of similar social media content (Del Vicario et al. 2017; Kim and Dennis 2019;

Moravec et al. 2019; Mutahi and Kimari 2020; Zollo 2019), few studies have considered the perceived similarity with an actor as a predictor of news believability. However, following the propositions of the similarity-attraction paradigm and the potential of social media to induce confirmation bias with similar others, we hypothesize:

Hypothesis 1: *Greater perceived similarity with an actor increases the believability of news compared to a dissimilar actor.*

Impact of Opinion Leaders on the Believability of False News

Social media also expand opportunities for actors to engage more broadly in immediate exchanges with users (Larson and Watson 2011; McFarland and Ployhart 2015; Treem and Leonardi 2012). Regular users, political advocates or (alternative) news media can freely create and share news without having to rely on gatekeeping traditional media (Linvill et al. 2019; Linvill and Warren 2020). Two-step flow theory helps to explain the impact that different types of actors (i.e., users, advocates, news media) have on the believability of news. It emphasizes the importance of opinion leaders, such as political advocates, in mass communication (Katz 1957; Katz and Lazarsfeld 1966). Opinion leaders are a particularly impactful minority (e.g., politicians, activists, influencers) with certain characteristics (e.g., strong political position, high social status) that enable them to interpret and disseminate news among regular users (Borge Bravo and Esteve Del Valle 2017; Dubois and Gaffney 2014). Their prominent position in the communication process gives opinion leaders a special power of persuasion among their followers (Weeks et al. 2017).

On social media, opinion leaders in form of purposive advocates (e.g., actors, politicians) (Guo et al. 2020; Guo and Zhang 2020; Sterrett et al. 2019; Wang et al. 2020; Zimmer and Reich 2018) have been shown to be particularly effective in sharing information and in triggering heterogeneous political discussions on different occasions. To the best of our knowledge, no study has yet conducted a comparative study on the distinct impact of different types of actors (i.e. users, politicians, news media) on the believability of mews on social media. Considering to the prominent role that politicians as opinion leaders have according to the two-step flow theory, we assume that similar advocates will have a stronger confirmation bias effect than news media or regular users:

Hypothesis 2: *Opinion leaders like politicians exacerbate the confirmation bias effects of perceived similarity on the believability of news compared to news media and regular users.*

Messenger Effect on the Believability of False News

The mediated communication process on social media allows to clearly delineate between the actor who creates information (i.e., source) and the actor who redistributes it (i.e., messenger). The literature on persuasion and attitudinal change, as well as behavioural economics, supports the notion of a "messenger effect" (Maclean et al. 2019). It distinguishes between the source as the actual creator of the

message (Ruggiero 2015) and the messenger as an agent who delivers information to the recipient (Dolan et al. 2012). The messenger effect describes that the value of an information strongly depends on the recipient's reaction to the messenger of that information (Hafner et al. 2019). Kassin (1983) demonstrated the messenger effect in a deposition testimony scenario where members of a jury were particularly influenced by the positive or negative demeanor of a surrogate witness (an irrelevant cue), who was simply reading a testimony on behalf of an absentee witness. Messengers have been found to influence the believability of information, especially when they share characteristics with the recipient (Maclean et al. 2019). To understand the confirmation bias effects of news, it is important to consider the differences between the source and the messenger of an item of information.

Most social media research leaves the option of an additional visible source uninvestigated and only investigates the messenger effect in combination with the characteristics of the article itself (Janze and Risius 2017; Kim and Dennis 2019; Kim et al. 2019). Evidence suggests that the messenger effect generally does not exceed the influence of the message itself in general (Austin and Dong 1994; Buchanan 2020; Dumitru 2020; Hafner et al. 2019; Joslyn and Haider-Markel 2006). However, when comparing the effects of the source and the messenger on social media, research shows that trust in the messenger more decisively influences engagement with news than the source (Sterrett et al. 2019). Further, information from a messenger sharing a distrusted media source who is a closer friend was perceived as more credible than information from a distant friend sharing a reliable news source (Oeldorf-Hirsch and DeVoss 2019). While no research has yet compared the simultaneous effects of messengers and sources on the believability of news on social media, given the evidence for the prevalence of a messenger effect, we follow the persuasion and behavioral economics literature and conclude:

Hypothesis 3: *The messenger exacerbates the confirmation bias effects of perceived similarity on the believability of news compared to the source.*

Impact of the Mediated Communication on the Believability of False News

Social influence describes the many ways in which people affect each other regarding, including changes in attitudes, beliefs, feelings, and behaviors as a result of comments, actions, or simply their presence (Sherif 1936; Sherif 1961). Social influence is exerted through conformity to group norms (Gilovich et al. 2010). In his seminal conformity experiments, Asch (1951) was able to include participants to give a clearly wrong answer in a simple perceptual task (i.e., determining which of three lines was the same length as a target line), by having confederate participants unanimously agree on the clearly wrong choice. These normative conformity pressures are particularly strong in the case of greater perceived similarity with the other actors (Rimal et al. 2005), at least to a size of three or four members (Gilovich et al. 2010). Having several similar people agree on a particular issue should create conformity pressures that bring people's beliefs in line with the group norms.

Social media platforms can increase normative conformity pressures by allowing multiple similar actors to share or retweet news. Redistributing others' posts and leveraging the source's reputation increases normative pressures compared to an individual actor who is solely dependent on their own perception when creating an original post. To the best of our knowledge, no research has yet considered the role of mediation on news believability. However, following the social influence theory, we assume that these conformity pressures to believe a particular article increase with the number of similar actors sharing it (Gilovich et al. 2010), e.g., in the case of mediated as opposed to immediate social media communication. This assumption received initial support by several studies, which show that the repeated exposure to even extremely implausible news stories can increase their believability (Pennycook et al. 2018; Shao et al. 2018a). In summary, we argue that a larger number of similar actors sharing news increases the probability that recipients will believe the news:

Hypothesis 4: *Mediated communication exacerbates the confirmation bias effects of perceived similarity on the believability of news compared to immediate communication.*

Further Theoretical Control Variables

Prior related work on the on the similarity-attraction paradigm and on news believability, proposes the relevance of additional variables. In particular, social desirability, political bias, reputation and liking (Roth et al. 2019; Wade et al. 2020), as well as the personal topical bias and personal relevance of the issue described in the article (Kim and Dennis 2019), and political affiliation as Democrat or Republican (Axt et al. 2020; Grinberg et al. 2019a; Guess et al. 2019) accompany the effects of news stories and the implications of similarity. We, therefore, include these effects as control variables in our empirical analysis.

Research Method

Experimental Setup and Manipulations

To test our hypotheses, we designed two comparable online experiments for the immediate (Study 1) and mediated communication process (Study 2) on Facebook (Table 2). In Study 1, we conducted a 2x3 mixed factorial experiment with the between-subject factor political bias (i.e., similar, dissimilar) and within-subject actor type (i.e., news media, politician, user). In Study 2, we extended this to a 2x2x3 design by adding the within-subject factor communication role (i.e., source, messenger). Lastly, we extended Study 2 to account for the cases where users receive news from similar and dissimilar actors (Bakshy et al. 2015; O'Hara and Stevens 2015) either as source or messenger.

To this end, we created Facebook profiles (Figure 5) representing different types of actors (i.e., news media, politician, user) with different political biases (i.e., similar, dissimilar) relative to the participant (i.e., Democrat, Republican) and manipulate the ways in which they share news (Figure 6 and Figure 7)

in terms of different communication processes (i.e., immediate, mediated) and in different communication roles (i.e., messenger, source).

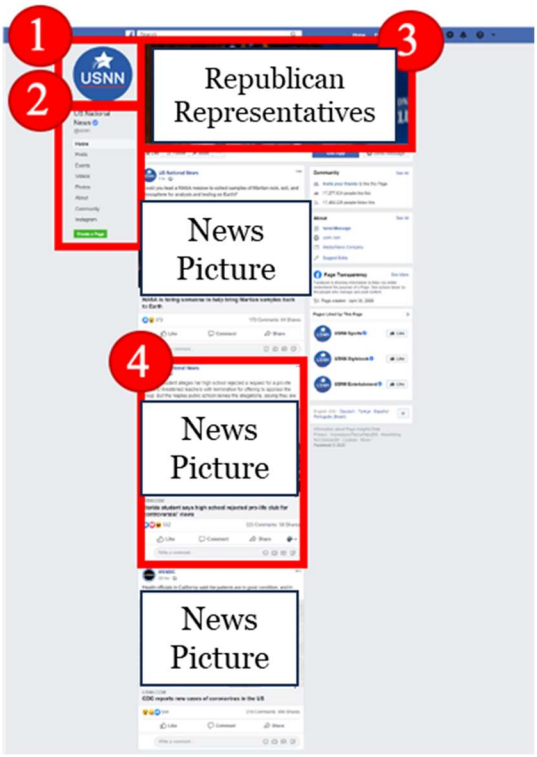
Example Actor: News Media	Manipulations
	**Actor Type** 1. Avatar indicates whether page is run by a news organization or a user (politician, user). 2. Facebook site: A Facebook *page* suggests that this page is run by a news organization or a political representative, and Facebook *profile* presents a private individual (user) **Political Affiliation** 3. Facebook cover image resembles a prominent political figure (associated with the party) for gatekeepers or party symbolism (donkey/elephant) for advocates and users 4. News article which resembles a traditional Democratic or Republican stance and a positive post expressing support for the Democratic (pro-gun safety) or Republican position (pro-life)

Figure 5. Exemplary manipulations of actor characteristics.

For both studies, we created Facebook sites to manipulate the actor types (Table 8, Table 9, and Table 10 in Appendix A). Facebook allows for the creation of sites, that can be either "pages" for commercial purposes or "profiles" (e.g., for individual social media users) with obvious structural differences. For each actor type, we chose the appropriate Facebook page or profile. We used Facebook "pages" for news media and politicians, and "profiles" for users to create our manipulations (Figure 5). In our study, we manipulated the avatar to represent either portray a news organization's logo (United States National News), a politician (Representative John Shane) with a substantial follower base and proposed policy or a regular user (Brian Smith) who resembled a private citizen and included information about hobbies and friends.

To manipulate the political bias (similar, dissimilar), we modified the content on the Facebook sites to include visible clues of Democratic or Republican Party affiliation. We drew on related studies to manipulate political similarity (Roth et al. 2019; Wade et al. 2020); each account site included cues that highlighted Democratic or Republican Party affiliation. We used clear political symbolism, i.e., the donkey (for Democrats) and the elephant (for Republicans), as well as topics that are stereotypically associated with Democrats (e.g., pro-gun control) and Republicans (e.g., pro-life) (Table 8, Table 9, and Table 10 in Appendix A).

We first used a cover image that clearly expresses party affiliation, in combination with the Democrat or Republican party symbol for advocates and users. Furthermore, each profile contained posts in clear support of either Republican or Democratic activities (e.g., "I'm voting Republican" or "We stand with Dreamers"). Political similarity or dissimilarity were established by exposing users to either similar or dissimilar actors depending on their own personal party affiliation (see also study process).

Immediate communication
Example: Politician communicates news article

Figure 6. Illustration of an immediate communication process on social media

Mediated communication
Example: User (messenger) shares news article from news media (source)

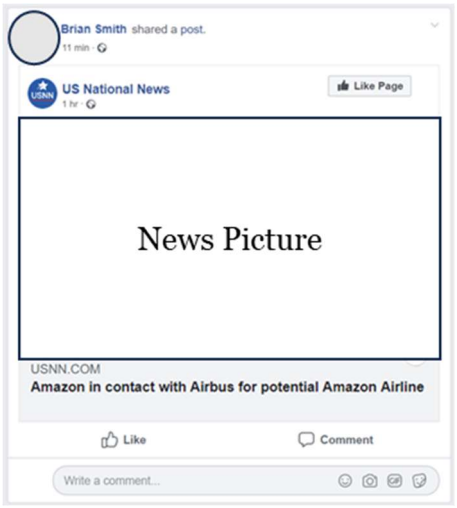


Figure 7. Illustration of a mediated communication process on social media

For the accounts, we reviewed real-world examples of each account type to establish ecological validity. We tailored the accounts accordingly by incorporating original content based on our review, and used generated names, avatars, and complementary information to ensure anonymity in this procedure. To further establish authenticity, we added information such as shares, likes, and nonrelated posts based on based on suggestions from the literature on opinion leaders (e.g., news organizations, politicians with high political power, large follower bases) and regular users (e.g., personal accounts, smaller follower bases) (Borge Bravo and Esteve Del Valle 2017; Dubois and Gaffney 2014).

Furthermore, we manipulate the communication process (Study 1: immediate, Study 2: mediated) by exposing participants to news that are either directly posted by an actor (Figure 6) or redistributed as shared by another actor (Figure 7). As explained above, in the commonly assessed immediate condition messenger and source are the same. However, in the commonly disregarded mediated condition, we can disentangle the actor roles by distinguishing between the original creator of a post (source) and the actor who shares the post (messenger).

We fabricated 12 news headlines (Table 1) based on examples from the literature (Moravec et al. 2019), choosing stories that are not directly associated with either a Democratic or a Republican position and that were unlikely to elicit a specific opinion or push the reader in a certain direction (full overview in Table 11 in Appendix B). This allowed us to focus solely on the similarity of accounts, their type, and the communication role. To control for influences of the post itself (Janze and Risius 2017; Vosoughi et al. 2018), we refrained from using accompanying text in the post. This ensures that the observed variance in news believability can be directly related to the actual experimental manipulations of perceived actor similarity, actor type, communication process, and communication roles.

#	Fabricated news stories
1	ISIS Leader Calls for American Muslims to Support Women's March.
2	China to Discontinue all Trade with the United States.
3	Amazon in Contact with Airbus for Potential Amazon Airline.
4	The U.K. Is about to Cut all Trade with the E.U., Focusing on the U.S. and Canada Instead.
5	Animal Migration Doors to be Installed in Border Wall to Appease EPA.
6	The United States Is Suing Volkswagen Over Cheating on Environmental Regulations.
7	Alphabet (Google) Is Being Sued for Record Sum for Mishandling Customer Data.
8	November Unemployment Rate Unexpectedly Rises to 6%.
9	Apple and Samsung Under Investigation Concerning Price Rigging in the Smartphone Market.
10	Daimler Plans to Shut Down Fabrication in the U.S.
11	Strategic Decision: Ford Aims to Step Back from Electric Mobility.
12	Microsoft Is Considering Moving Production out of China.

Table 1. Descriptive news headlines, not directly associated with a Democratic or Republican stance.

Study Process

Both studies follow a three-stage process of pre-, main-, and post-study (Figure 8) with accompanying measurements (described below in Operationalization and Measurements).

During the pre-study, we assess the participants' political affiliation as either Democrat or Republican before evenly assigning them to the political (dis)similarity bias. Participants who do not report a political bias that is aligned with either one of these two parties are excluded from participation. Participants are then assigned to either the direct or mediated study.

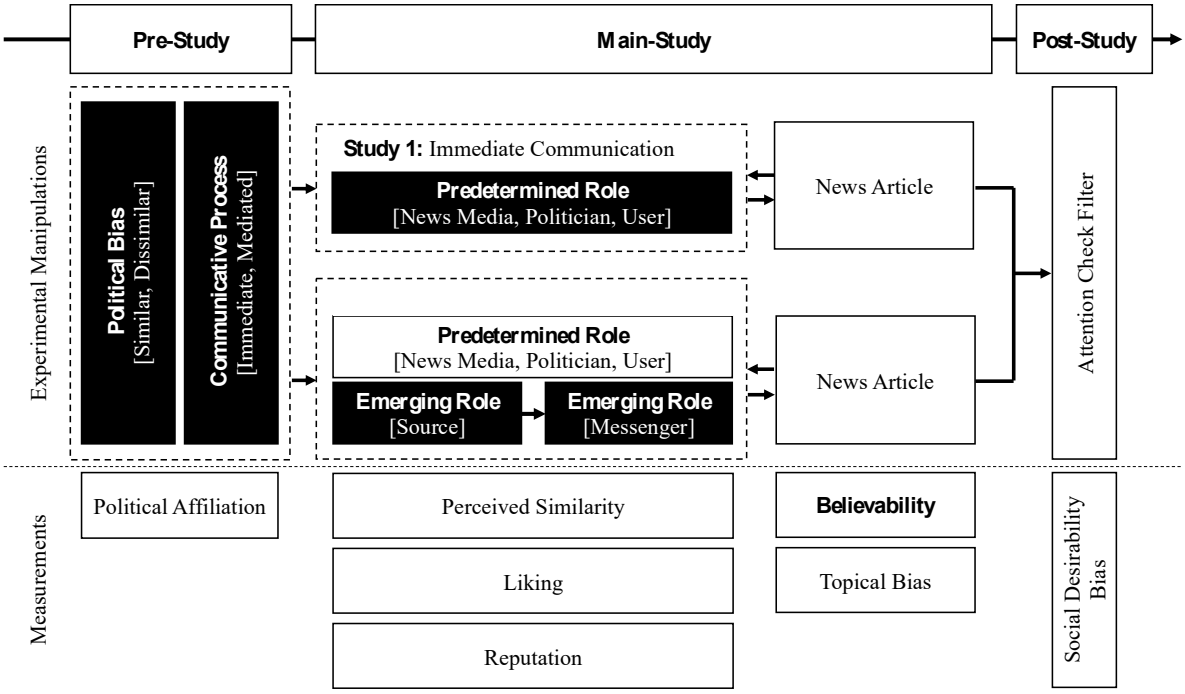


Figure 8. Study process with experimental manipulations and measurements.

In the main study, participants are first exposed to the Facebook site of one of the types of actors (news media, politician, or user) and rate their perceived similarity, liking and reputation along with questions related to the politically charged content in the actor's timeline to ensure that participants paid attention and processed the manipulations appropriately. In the immediate condition (study 1), participants then assess four randomly drawn posts of that actor regarding their perceived believability and the topical bias before repeating the process for the other two actors in a random sequence. In the mediated condition (Study 2), participants are initially exposed to the Facebook sites of two actors before processing six mediated posts with messenger and source held constant. Subsequently, the participants are presented with the third actor site before rating the second set of six mediated posts with the new actor as the messenger. Throughout the study at the first, the sixth, and the ninth news article, participants are also asked to check a particular answer category as an attention check. After processing all twelve news stories, in the post study respondents answer questions about social desirability before being debriefed and asked for feedback on the study.

Study 1: Immediate Communication Process				
Actor Type		G	A	U
Political Bias	Similar	G→	A→	U→
	Dissimilar	G-→	A-→	U-→

Study 2: Mediated Communication Process										
Political Bias	**Source**	**Messenger**	→							
Similar	U→	G→	A→	G→	G→	A→	U→	A→	G→	U→	A→	U→
Dissimilar	U-→	G-→	A-→	G-→	G-→	A-→	U-→	A-→	G-→	U-→	A-→	U-→
Dissimilar Gatekeeper	U→	G-→	A→	G-→	G-→	A→	U→	A→	G-→	U→	A→	U→
Dissimilar Advocate	U→	G→	A-→	G→	G→	A-→	U→	A-→	G→	U→	A-→	U→
Dissimilar User	U-→	G→	A→	G→	G→	A→	U-→	A→	G→	U-→	A→	U-→

Notes.→ similar, -→ dissimilar; Gatekeeper (G) = News Media, Advocate (A) = Politician, User (U); 30 Democratic and Republican participants per cell; $N_{Immediate}$ = 120 participants, $N_{Mediated}$ = 900 participants; grey = similar /dissimilar only cases for hypothesis 4 testing

Table 2. Overview of study 1 and 2 with immediate and mediated conditions.

The studies took an average of 13 minutes to complete and were posted on surveymonkey.com. We recruited the sample on Amazon Mechanical Turk (mTurk), and each participant was compensated with 1.00 USD. Regarding mTurk, there is a risk that we cannot know whether the participants are authentic. We followed the suggestions of Lowry et al. (2016) who outline several recruiting characteristics to counteract such risk. Accordingly, we recruited only mTurk participants who had (1) earned the master's qualification, (2) demonstrated high approval ratings and (3) a history of overall satisfactory task performance. Furthermore, the location of the participants was set to the United States. We followed common guidelines for the number of observations for the following hierarchical regressions (Hox 2010): small groups on level 1 with 5-10 observations require at least 30-50 individuals on level 2. Given that we had 12 level 1 observations per participant, we chose 30 participants per political affiliation (i.e., Democrat or Republican) totaling 60 participants per condition. Accordingly, we recruited a sample of 120 participant in study 1 and 900 participants in study 2 (with four randomly missing observations due to nonresponses and one completely withdrawn participant).

Operationalization and Measurements

We used instruments that were all validated by prior research to gather information about the participants themselves, their perception of the actors and their assessment of the news stories.

For the article specific measures, we assess the dependent variable believability through the average of three items on a seven-point scale ranging from "Strongly Disagree" to "Strongly Agree" to see whether each article was perceived as believable, credible and truthful (Kim and Dennis 2019). We also assess the topical bias of an individual towards each article in terms of the importance of the issue described in the article ("Do you find the issue described in the article important" 1 = not at all, 7 = extremely) multiplied by the personal position regarding the topic of the article ("What is your personal position towards the article?" -3 = extremely negative, +3 = extremely positive) (Kim and Dennis 2019).

To evenly assign participants to the political bias treatment groups, each participant is asked about their political party affiliation at the beginning of the experiment. They could choose "Democrat", "Republican" or "other." Participants who chose "other" are disqualified from the experiment, while "Democrats" and "Republicans" are assigned to the experimental conditions. The specific political affiliation is later used as a control variable since individual findings indicate differences between Democrats and Republicans regarding their susceptibility to belief in deceptive information (Axt et al. 2020; Grinberg et al. 2019a; Guess et al. 2019). Also, we used an established measure of social desirability bias (Reynolds 1982). The 10 items are measured on a seven-point scale ranging from "Strongly Disagree" to "Strongly Agree".

Regarding the relationship with the actor, we first assess the perceived similarity as the study's key independent variable and predictor of the similarity-attraction paradigm (Roth et al. 2019; Wade et al. 2020). To that end, we use a seven-point scale ranging from "Strongly Disagree" to "Strongly Agree", adapting the instrument of Tepper et al. (2011) to the context and the specific actor. We also assess additional control variables that have been shown as impactful in the similarity-attraction paradigm context (Roth et al. 2019; Wade et al. 2020). We measure the liking of the participant for the actor with the scale from Wayne and Ferris (1990) adapted to the context (person or Facebook page) on a five-point scale ranging from "1 = I don't like this person/site at all" to "5 = I like this person/site very much." Lastly, we measure the actor's perceived reputation by asking whether or not they trusted the actor in line with (Kim and Dennis 2019). Participants answer "[Actor] is a trusted news source" on a seven-point scale ranging from "Strongly Disagree" to "Strongly Agree."

Empirical Results

Descriptive Statistics and Manipulation Check

To test the hypotheses, we used multilevel fixed-effects regressions with restricted maximum likelihood estimates (Snijders and Bosker 2012). In our regression models, we accounted for effects of the individual news stories, the study participants themselves, and the actors that we manipulated as part of the experiment. In particular, we controlled for the different levels of believability of the individual news stories by assigning and including a unique identifier for each article (article type number 1-12) as a categorical variable and the topical bias expressed by the participants for each article. This helped ensure that we test the confirmation bias effects of the actor characteristics instead of the previously addressed effects of the message content (Kim and Dennis 2019). In addition, we included the participants' self-reported social desirability and their political affiliation (i.e., Democrat, Republican). Lastly, we included the participants' score of their assessed liking and reputation (Kim and Dennis 2019) of the displayed actors (i.e., news media, politician, user).

We used two-level hierarchical models where the 12 individual observations of the stories and their believability assessments (Level 1) were nested in the individual participants (Level 2). This allowed us to leverage the repeated measurements for each participant (Toutenburg and Shalabh 2009) and further

ensure that the observed effects were related to the experimental manipulations of the account characteristics. The intraclass correlation coefficients confirm that the proposed nested structure as the reliability of participant means is expectably low for each of the following regressions (ICC1 = 0.1891, ICC2 = 0.2129, ICC3 = 0.2293) (Bliese 2000). Additionally, for each model we compared the mixed model with a standard OLS regression (Hox 2010). The results continuously show a significant increase in model fit ($p < 0.001$) for the hierarchical structure compared to the standard OLS regression. Accordingly, we accepted the multilevel structure and evaluated the model by comparing their goodness-of-fit test statistics (McCullagh and Nelder 1989). Specifically, we computed the established restricted maximum log-likelihood values, Akaike's information criterion (AIC) (Box et al. 2007), and the Bayesian information criterion (BIC). Looking at the goodness-of-fit measures, the decreased model fit scores compared to the intercept-only and control variable models respectively confirm the validity and appropriate complexity of our proposed model. Furthermore, the computed pseudo R2 scores (R2 = 1-(Vfull / Vintercept only); R21 = 36.32%, R22 = 34.44%, R23 = 36.03%) indicate an acceptable amount of explained variances for all models.

Political affiliation	Immediate			Mediated	
	G	A	U	Consistent	Mixed
Believability					
Similar	3.72	3.67	4.21	3.9	3.88
Dissimilar	3.64	3.4	4.06	3	3.89
Perceived Similarity					
Similar	4.33	4.76	4.58	4.55	4.24
Dissimilar	4.11	3.86	3.06	3.71	4.15

Notes. $N_{immediate}$ = 120 participants, $N_{mediated}$ = 900 participants; G = News media (gatekeeper), A = Politician (advocate), U = User, "Mixed" condition from messenger perspective (i.e., similar or dissimilar messenger, source is respectively reversed)

Table 3. Descriptive statistics for the treatment effects on news believability and perceived similarity.

The inspection of the descriptive statistics (Table 3Table) confirms the effectiveness of our manipulations and supports the assumed confirmation bias effect. For the immediate condition (Study 1), the average believability was higher for similar accounts ($\bar{x}_{SimilarGkp}$=3.72, $\bar{x}_{SimilarA}$=3.67, $\bar{x}_{SimilarU}$=4.21) compared to their dissimilar counterparts ($\bar{x}_{DissimilarGkp}$=3.64, $\bar{x}_{DissimilarA}$=3.4, $\bar{x}_{DissimilarU}$=4.06). Simultaneously, the perception of similarity was also higher for similar accounts ($\bar{x}_{SimilarGkp}$=4.33, $\bar{x}_{SimilarA}$=4.76, $\bar{x}_{SimilarU}$=4.58) compared to the dissimilar accounts ($\bar{x}_{DissimilarGkp}$=4.11, $\bar{x}_{DissimilarA}$=3.86, $\bar{x}_{DissimilarU}$=3.06). In the mediated experimental design (Study 2), the condition with consistently similar political actors was more believable ($\bar{x}_{Similar}$=3.9, $\bar{x}_{Dissimilar}$=3) and seemed more similar ($\bar{x}_{Similar}$=4.55, $\bar{x}_{Dissimilar}$=3.71) than the concatenation of dissimilar actors. As expected, these global differences were not as apparent in the mixed condition with similar messengers and dissimilar sources or vice versa for believability ($\bar{x}_{Similar}$=3.88, $\bar{x}_{Dissimilar}$=3.89) and similarity ($\bar{x}_{Similar}$=4.24, $\bar{x}_{Dissimilar}$=4.15), which allowed us to used that condition to disentangle messenger and source effects in the following.

Inferential Statistics and Hypothesis Testing

To test the effects of the hypothesized individual fixed effects (Figure 4), we calculated the corresponding Satterthwaite-corrected *t*-statistics and *p*-values for the path coefficients (Hox 2010). The hypothesized moderator effects evaluated using the combined consideration of main and interaction effects (Baron and Kenny 1986). As denoted in the individual regression models, we performed omnibus tests in the form of simple contrasts of marginal linear predictions for multicategorical dummy variables and Wald tests for the interaction terms with categorical variables where applicable (Wooldridge 2010).

Fixed Effects	**Coefficient**	**S.E.**	**Coefficient**	**S.E.**
Intercept	2.327***	.294	2.6***	.32
Controls				
Article type[a]	$\chi^2_{(11)} = 406.1^{***}$		$\chi^2_{(11)} = 397.55^{***}$	
Article topical bias	.067***	.005	.067***	.005
Participant social desirability	.05*	.06	.053	.06
Participant political affiliation	-.209†	.125	-.208†	.126
Actor liking	.033	.057	.09	.059
Actor reputation	.121**	.036	.111**	.037
Hypotheses				
Perceived Similarity	.078†	.041	.009	.056
Actor Type (reference group: News media)[a]			$\chi^2_{(2)} = 9.66^{**}$	
Politician			-.719**	.232
User			-.371†	.221
Perceived Similarity x Actor Type[b]			$\chi^2_{(2)} = 4.68^{\dagger}$	
Politician			.108*	.053
User			.035	.052
Random part	Var. comp.	S.E.	Var. comp.	S.E.
Level-two variance: Participants	.3286	.0603	.337	.062
Level-one variance: Observations	1.4595	.057	1.445	.057
Log-likelihood (ML)[b]	-2,434.38***		-2,434.16***	
AIC	4,908.476		4,916.324	
BIC	5,014.09		5,043.06	
Pseudo R^2			.3632	

Notes. Dummy coded variables (with reference category): political affiliation (Democrats), actor type (news media), article type (first article); 12 article types included in calculation but only summative test score depicted for clarity, x = Interaction term, ML = maximum likelihood, AIC = Akaike information criterion, BIC = Bayesian information criterion.
Sample. Participants (level two) = 120, Individual observation (level one) = 1,440.
p-values. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.09$ (two-tailed significance).
Parameter tests: [a]simple contrasts of marginal linear predictions: $\chi^2_{(df)}$, [b]Wald test: $\chi^2_{(df)}$

Table 4. Multilevel regression analysis testing the immediate effects of perceived similarity and actor types on news believability.

We first tested the effects of perceived similarity (*hypothesis 1*) and the enhancing role different actors (*hypothesis 2*) in the sample of immediate communication (Table 4). The results generally confirm a positive relationship between perceived similarity and news believability (French et al. 2023), albeit the relationship being a lot more intricate than commonly assumed. In the model without the actors, we find a marginally significant positive relationship between perceived similarity and news (b_{PS} = .078, p = 0.056). This positive relationship is fully qualified by the addition of actor types (b_{PS} = .009, p = 0.17) that shows a significant main effect of the actors ($\chi^2_{(2)}$ = 9.66, p = 0.008) and their moderation ($\chi^2_{(2)}$ = 4.68, p = 0.096). It is worth noting that inspecting the predictive margins (Figure 9), we find that news

from media organizations are generally more believable than news from politicians ($b_{Politician}$ = -.719, p = .002) and users (b_{User} = -.371, p = .094). Furthermore, the believability of news from politicians depends significantly more on their perceived similarity compared to news media ($b_{I\_Politician}$ = .108, p = .04).

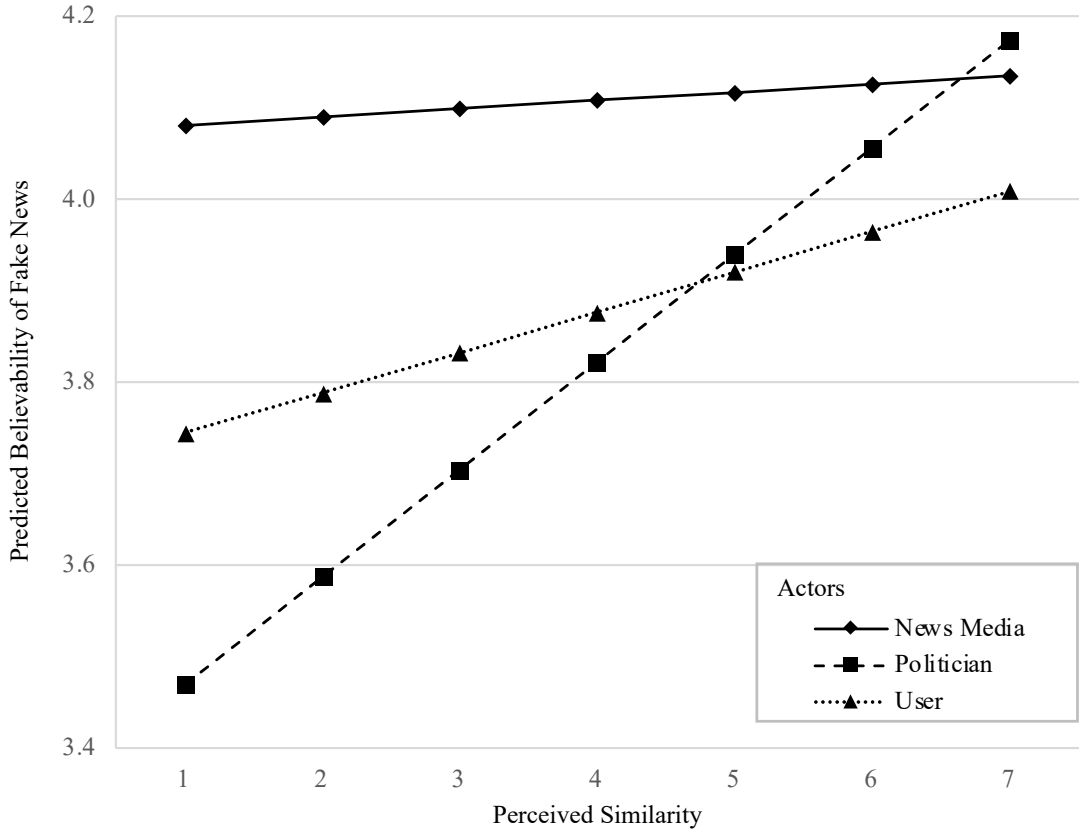


Figure 9. Interaction effects of actors' predetermined roles (i.e., news media, politician, user) with perceived similarity on news believability.

We then tested the differential effects of the messenger and the source (*hypothesis 3*) in the sample of mediated communication (Table 5). The results generally confirm a positive impact of the perceived similarity of the messenger ($b_{Messenger}$ = .117, p = 0.009). Different to what we had expected, the effect of the messenger does not overpower the effect of the sources' perceived similiarity that remains significant (b_{Source} = .092, p = 0.013). Instead, we find a far more intricate interaction between the perceived similarity of the source and messenger on the believability of news ($b_{I\_MessengerSource}$ = -.018, p = 0.012). The orthogonal interaction in the predictive margins illustrates an amplification effect (Figure 10). The greatest believability of a news article follows from the combination of a similar communicator sharing news from a dissimilar source. This emergent "mixed-actor-amplification-effect" goes beyond a mere confirmation bias where a similar actor would outweigh a dissimilar one (French et al. 2023). Instead, the significant negative interaction demonstrates that news believability is exacerbated when a similar and a dissimilar actor share the same news.

Fixed Effects	Coefficient	S.E.	Coefficient	S.E.
Intercept	2.616***	.179	1.979***	.221
Controls				
Article type[a]	$\chi^2_{(11)}$ = 1,567.55***		$\chi^2_{(11)}$ = 1,564.42***	
Article topical bias	.076***	.002	.076***	.002
Participant social desirability	.032*	.028	.03	.028
Participant political affiliation	.072	.058	.071	.058
Messenger liking	.073*	.035	.04	.045
Messenger reputation	.059*	.024	.047†	.025
Source liking	.063**	.022	.057*	.028
Source reputation	.088***	.017	.089***	.018
Hypotheses				
Perceived Similarity Messenger			.117**	.045
Perceived Similarity Source			.092*	.037
Perceived Similarity Messenger x Source[b]			-.018*	.007
Random part	Var. comp.	S.E.	Var. comp.	S.E.
Level-two variance: Participants	.361	.029	.355	.028
Level-one variance: Observations	1.313	.024	1.313	.024
Log-likelihood (ML)[b]	-10,869.854***		-10,875.386***	
AIC	21,781.71		21,798.77	
BIC	21,924.73		21,962.22	
Pseudo R^2			.3444	

Notes. Dummy coded variables (with reference category): political affiliation (Democrats), article type (first article); 12 article types included in calculation but only summative test score depicted for clarity, x = Interaction term, ML = maximum likelihood, AIC = Akaike information criterion, BIC = Bayesian information criterion.
Sample: Participants (level two) = 539, Individual observation (level one) = 6,464.
p-values. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, †$p < 0.09$ (two-tailed significance).
Parameter tests. [a]simple contrasts of marginal linear predictions: $\chi^2_{(df)}$, [b]Wald test: $\chi^2_{(df)}$

Table 5. Multilevel regression analysis testing the differential effects of the perceived similarity of the messenger and sources on news believability.

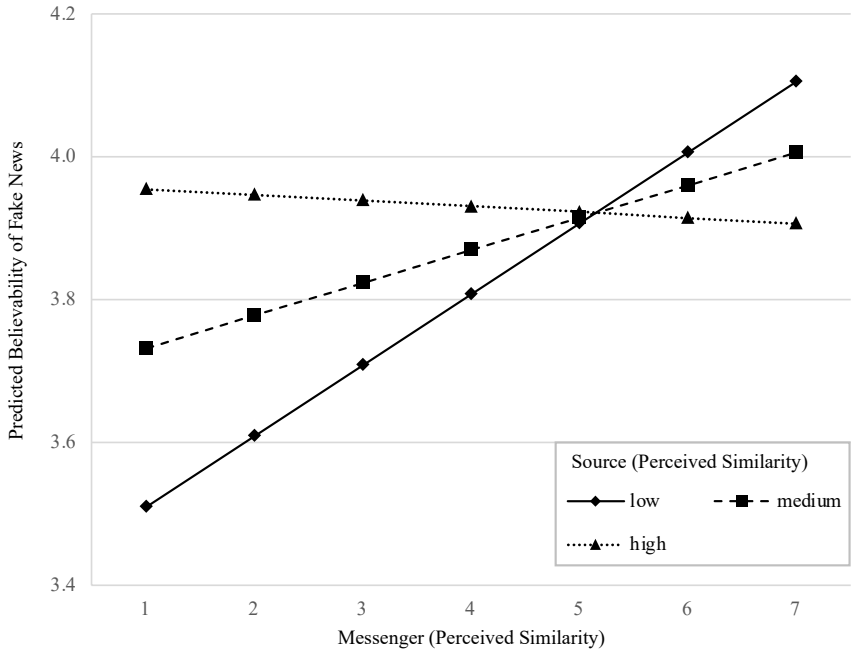


Figure 10. Interaction effects of the perceived similarity with actors of different emerging roles (i.e., messenger, source) on news believability.

Lastly, we test the effects of the mediated vs. immediate communication process and whether the concatenation of actors boosts the believability of news (*hypothesis 4*). Specifically, we compared the effects of purely (dis)similar conditions between the immediate and mediated samples (Tabel 6). In support of hypothesis 4, we find that the main effect of perceived similarity (b_{PS} = .049, p < .000) is significantly exacerbated through the mediated communication process when multiple (dis)similar actors share the same news ($b_{I\_PSCommunication}$ = .074**, p = .002). In other words (Figure 11), the believability of news increases (decreases) when two similar (dissimilar) messengers and sources share the same news as opposed to only one of them posting the news (immediate communication).

Fixed Effects	**Coefficient**	**S.E.**	**Coefficient**	**S.E.**
Intercept	2.205***	.11	2.368***	.118
Controls				
Article type[a]	$\chi^2_{(11)}$ = 2,686.94***		$\chi^2_{(11)}$ = 2,679.69***	
Article topical bias	.073***	.002	.072***	.002
Actor liking	.062**	.019	.088***	.045
Actor reputation	.138***	.015	.145***	.015
Participant social desirability	.043*	.02	.044*	.02
Participant political affiliation	-.004	.044	-.003	.044
Hypotheses				
Perceived Similarity	.049***	.011	-.002	.018
Communication (reference group: immediate)[a]			-.385***	.103
Perceived Similarity x Communication[b]			.074**	.024
Random part	Var. comp.	SE.	Var. comp.	SE.
Level-two variance: Participants	.3576	.022	.354	.022
Level-one variance: Observations	1.3384	.02	1.338	.02
Log-likelihood (ML)[b]	-16,829.03***		-16,826.805***	
AIC	33,698.06		33,697.61	
BIC	33,842.87		33,856.9	
Pseudo R^2			.3603	

Notes. Dummy coded variables (with reference category): political affiliation (Democrats), communication process (immediate), article type (first article); 12 article types included in calculation but only summative test score depicted for clarity, x = Interaction term, ML = maximum likelihood, AIC = Akaike information criterion, BIC = Bayesian information criterion.
Sample: Participants (level two) = 839, Individual observation (level one) = 7,914.
p-values. $^{***}p$ < 0.001, $^{**}p$ < 0.01, $^{*}p$ < 0.05, $^{\dagger}p$ < 0.09 (two-tailed significance).
Parameter tests. [a]simple contrasts of marginal linear predictions: $\chi^2_{(df)}$, [b]Wald test: $\chi^2_{(df)}$

Table 6. Multilevel regression analysis testing the complementary effects of actors in an (im)mediate communication process on news believability.

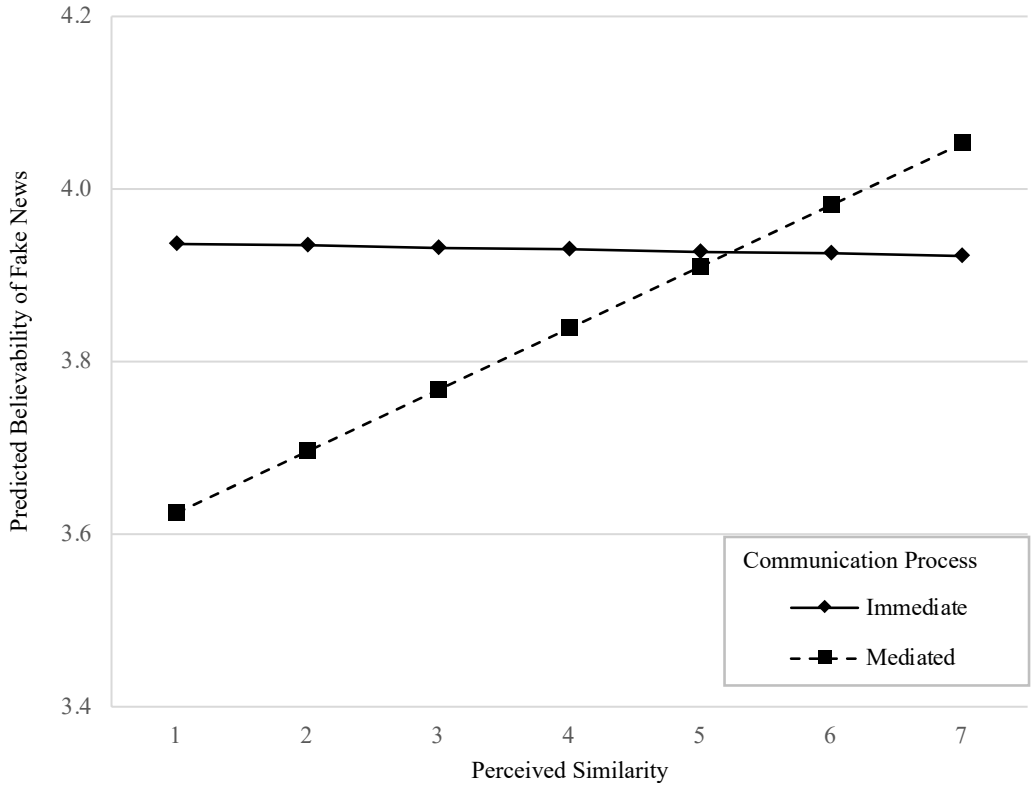


Figure 11. Interaction effects of the communication process (i.e., mediated, immediate) with perceived similarity on news believability.

Discussion

Motivated by a desire to understand how the disintermediation of communication by social media, limits institutional press' ability to engage with audiences and creates opportunities for new actors to shape news consumption, we investigated *how social media communication influences the believability of news*. To answer this research question, we drew on the traditional mass communication model (Westley and MacLean Jr 1957), to highlight and conceptualize how social media (1) liberated opportunities for competing actors (i.e., news media, politicians, users) to directly share news with the audience itself (Larson and Watson 2011), (2) encouraged proliferation of news through mediated redistribution of content (boyd et al. 2010), and (3) created new distinctions and emerging roles between messengers and sources in networked communication (Vaast et al. 2017). To evaluate our conceptualization, we performed two online experiments that test the impact of these social media communication characteristics on the believability of news along four hypotheses grounded in the similarity-attraction paradigm. Table 7 presents our hypotheses and results.

#	Hypothesis	Decision	Finding
1	*Greater perceived similarity with an actor increases the believability of news compared to a dissimilar actor.*	**(Partially) Supported**	While we generally find a positive relationship between perceived similarity with an actor and news believability, this relationship is subject to intricate moderation effects (see $H_{2\text{-}4}$).
2	*Opinion leaders like politicians exacerbate the confirmation bias effects of perceived similarity on the believability of news compared to news media and regular users.*	**Supported**	Politicians are significantly more affected by the confirmation bias effect of perceived similarity than news media. News media are generally more believable than users or politicians, irrespective of the perceived similarity.
3	*The messenger exacerbates the confirmation bias effects of perceived similarity on the believability of news compared to the source.*	**Not supported**	The perceived similarity of messengers and sources have independent main effects. We also observe a "mixed-actor-amplification-effect" whereby news communicated from one similar and one dissimilar actor appear most believable.
4	*Mediated communication exacerbates the confirmation bias effects of perceived similarity on the believability of news compared to immediate communication.*	**Supported**	Mediated communication of two concatenated (dis)similar actors boosts the (dis)confirmation bias effect of perceived similarity.

Table 7. Overview over the tested hypotheses and findings observed across two studies.

Our results suggest five key observations. First, while controlling for the topical bias of the news story (Kim and Dennis 2019) we investigate the characteristics of the actors themselves. We find that users' perceived similarity with the actor on social media affects the believability of news. People tend to believe news from others who have a similar outlook, perspective, and values. However, this relationship is not uniform but subject to the following moderating circumstances (i.e., actor types, mediated communication process, interplay between sources and messengers). Hence, while we find support for the general assumption that greater perceived similarity with an actor improves the believability of the news they share, this relationship is subject to moderating circumstances.

Second, we found that predetermined role of the actor on social media – as news media, politician, or user – affects news believability. Despite the growing distrust in the traditional press (Luo et al. 2022; Valdez and Ziefle 2018), information shared by news media appears to be generally more believable than those of partisan users or politicians – irrespective of their perceived (dis)similarity. For news media, we do not find a confirmation bias effect of the perceived similarity. This is different for opinion leaders like politicians, whose believability is moderated by their perceived similarity (Weeks et al. 2017). When a user feels more similar to a politician, they find news shared by them as more believable. Nonetheless, when compared to overtly partisan users or politicians, news media organizations appear to retain a greater credit of trust.

Third, we find evidence for a "mixed-actor-amplification-effect" where a similar messenger sharing the news of a dissimilar source can actually exacerbate the news believability. This demonstrates that the messenger cannot only dilute the effects of the source (Sterrett et al. 2019), but also reinforce their believability. This mixed-actor-amplification-effect is different from other decision-making influences

like a confirmation bias, arguing against self-interest, or representativeness heuristics (Gilovich et al. 2010). Existing biases focus on the characteristics of the message or messenger: is the content similar to the own view, does the messenger convey a message that hurts their personal interests, or does the messenger fulfill the assumed stereotypical characteristics of a certain group. The emergent mixed-actor-amplification-effect considers the combination of multiple actors and finds that news are particularly believable when they are shared by two actors from different camps.

Fourth, we did not find evidence for a superimposing effect for the perceived similarity with a messenger over the news source. Instead, both emerging roles have independent main effects. This emphasizes the importance of considering the particularities of social media communication where source and messenger are both distinctively apparent. The commonly accepted "messenger effect" was established in situations where the source was removed from the immediate conversation (Kassin 1983; Maclean et al. 2019). Hence, the characteristics of the messenger is able to determine the believability of a message. On social media, however, the source is immediately apparent next to the messenger. In these cases, the perceptions of the source and the messenger become relevant for the impact of news stories. As demonstrated by the aforementioned mixed-actor-amplification-effect, the interactions between these emerging roles can produce unexpected consequences.

Fifth, we find that concatenating two (dis)similar actors exacerbates the believability enhancing (diminishing) effects of the perceived (dis)similarity. A user or politician with a similar political affiliation sharing news from a politically similar news organizations boosts their believability. This helps explain the particular efficacy of echo chambers wherein members are not only selectively exposed to information (Kitchens et al. 2020), but the redistribution of information among like-minded people itself makes them more believable.

Theoretical and Practical Implications

In essence, this article demonstrates the importance of considering the unique communication characteristics of social media compared to traditional communication, separate cleanly between message and actor effects, and to accept more nuance in the conversation regarding the confirmation bias of (dis)similar actors. Thereby, this study makes several key practical and theoretical contributions.

This study provides practically relevant insights for news media organizations, legislators, and society at large. For news media, understanding the process of how social media communication affects their believability to guide their social media management efforts. Our findings suggest that the believability of news organizations is not significantly affected by how they are perceived themselves, but instead by who shares their news. In that regard, our study allows to derive two suggestions. First, news organizations ought to convince their followers to redistribute their stories – not just to expand the spread and reach but also – to boost their believability. Second, the mixed-actor-amplification-effect highlights the importance of also engaging with dissimilar actors. Ignoring dissimilar actors is a missed opportunity for news media to have a meaningful impact to other audiences. This exceeds the current conversation that focuses, for example, on the effects of journalists who engage directly with audiences

online (Lee 2015). For legislators, it is imperative to have a proper understanding of the social media communication process. As we outline in greater detail above, malicious actors abuse the information void created by institutional media's inability to reach and inform the public for propaganda purposes, disinformation campaigns, and societal polarization (e.g., by creating fake politician, user, and news media accounts or by paying-off influencers) (Linvill and Warren 2020; Ryan et al. 2022). Properly diagnosing the problem in terms of the social dimension of how people engage with information in the current information ecosystem, is the foundation to remedying it (Graham).

This line of research also has important theoretical implications, aside from filling the aforementioned knowledge gaps regarding the conceptualization of the particularities of social media communication or the confirmation bias effects of actor characteristics beyond message effects. We demonstrate the importance of distinguishing between the predetermined and emerging roles of actors on social media. Although much news research currently seems to assume that the source and messenger are the same (Kim and Dennis 2019; Kim et al. 2019), traditional communication research has established the importance of accounts used to source versus process information (Lee and Sundar 2013; Metzger et al. 2010; Wilson and Sherrell 1993). Since researchers have argued that opinion leaders do not form a structurally separate group from the social systems of regular users (Borge Bravo and Esteve Del Valle 2017), our findings underscore the importance of considering communication roles to understand how news disseminate (Vosoughi et al. 2018). Based on our findings, future research might consider how the interplay between account types (i.e. those of opinion leaders) and drivers of perceived similarity (e.g., political affiliation, lifestyle, or brand support) could impact the dissemination of content within an immediate community but also among interlinked online and offline communities.

Limitations and Future Research

A common concern about experiments is their external validity. Accordingly, we only used fabricated news and actors to strengthen the internal validity of our conclusions. While this helped to ensure our deductions were based on our experimental manipulations, the external validity could be strengthened, for example, by replicating these findings using real world actors and news (Starbird et al. 2023) or by exploring the effects of the message text that accompanies a news post (Janze and Risius 2017). Beyond complaints about external validity, complicated experiments such as ours are sometimes criticized for confounds that limit internal validity. To minimize the risk of carry-over effects, we randomized the sequences of fake news and accounts while maintaining the sequence of seeing the messenger first and then the source. This allowed us to differentiate between roles, but participants may have paid more attention to the first profile (messenger) than the second (source). Alternatively, participants may have become more familiar with the messenger because this role stayed the same across the 12 stories, whereas the source changed after six stories. However, the descriptive statistics suggest that the opposite is true when comparing originators and messengers and the participant's comments do not indicate any such bias. Nonetheless, further research could test for these possibilities by replicating the study with a balanced complete block design of experimental account characteristic conditions.

By outlining the expanded opportunities for mediated and immediate exchanges on social media, this study lays the foundation for a new research agenda on the role of the institutional press in the digital

age. For example, this study demonstrates the need to further investigate the interplay between emerging roles (e.g., source, messenger, external platform source) and predetermined actors (e.g., news media, users, social activists, influencers and content creators, companies) on the dissemination of information on social media. There are indications that celebrities as messengers have a particularly influencing effect over the source characteristics (Sterrett et al. 2019). For example, future research ought to consider the role of news media as platform-external originator of news, messenger or source of information. Furthermore, to understand messenger effects, we need to better understand the extent to which news are shared. For example, how many messengers share the content of a single source? Is it five or fifty or five hundred? And what effects do these numbers have on believability? This issue is complex because messengers may share news about the same issue but from different sources, as is common during global protests such as the Black Lives Matter movement. Understanding the opinion-forming effects of the sharing dynamics of social media and source-messenger combinations is essential to fully understand how mass online events develop (Drasch et al. 2015; Jöntgen 2020). Lastly, this study focuses on understanding how account characteristics moderate the relationship of confirmation bias and believability but leave unexamined the opinion-polarizing effect of news. The impact of opinion polarization on public discourse is considered a threat to modern society (Sunstein 2018). Polarization can lead to severe societal impacts such as terrorism or other violent events (Cheung et al. 2016; Hamm and Spaaij 2015; Wong et al. 2018). Alternatively, information diffusion theories like the spiral of silence describe more complex communication circumstances in terms of private and public opinions, the willingness to express them, and the impact of counterspeakers (Taylor 1982). It raises questions, for example, on what happens if a (dis)similar actor provides disconfirming information and how this can affect personal echo chambers or information cocoons. Future research could build on our findings to identify circumstances under which different accounts cause polarization and lead to harmful behaviors.

Conclusion

Social media platforms have disintermediated the information ecosystem. This transformation created an information void by diminishing the gatekeeping function of the institutional press between politicians and users. In this paper, we highlight three particularities of social media communication, namely broader opportunities to engage with audiences for all actors, the redistribution of news stories among actors, and the apparent distinction between messengers and sources. We then test how these social media communication characteristics influence the believability of news stories online. Our results demonstrate the role of biases found in other settings on news believability (i.e., opinion leadership, social influence), shows nuances to commonly accepted phenomena (i.e., similarity-attraction paradigm, messenger effect), and identifies newly emerging effects (i.e., mixed-actor-amplification-effect). Overall, our findings demonstrate the complexity that news media face when disseminating information, highlights the impact of social media communication characteristics, and allows to derive ways for news media to work towards a more positive future with an informed public.

References

Allcott, H., and Gentzkow, M. 2017. "Social Media and Fake News in the 2016 Election," (31:2), pp. 211–236.

Anspach, N. M. 2017. "The New Personal Influence: How Our Facebook Friends Influence the News We Read," *Political Communication* (34:4), pp. 590-606.

Asch, S. 1951. "Effects of Group Pressure Upon the Modification and Distortion of Judgments," in *Groups, Leadership and Men: Research in Human Relations,* G. Guetzkow (ed.). Pittsburgh, PA: Carnegie Press, pp. 177-190.

Aubin, C. S., and Liedke, J. 2024. "News Platform Fact Sheet." Retrieved November 11, 2024, from https://www.pewresearch.org/journalism/fact-sheet/news-platform-fact-sheet/#:~:text=News%20across%20digital%20platforms,-There%20are%20several&text=News%20websites%20or%20apps%20are,and%205%25%20who%20prefer%20podcasts.

Austin, E. W., and Dong, Q. 1994. "Source V. Content Effects on Judgments of News Believability," *Journalism Quarterly* (71:4), pp. 973-983.

Axt, J. R., Landau, M. J., and Kay, A. C. 2020. "The Psychological Appeal of Fake-News Attributions," *Psychological Science* (31:7), pp. 848-857.

Bakshy, E., Messing, S., and Adamic, L. 2015. "Exposure to Ideologically Diverse News and Opinion on Facebook," *Science* (348:6239), pp. 1130-1132.

Baron, R. M., and Kenny, D. A. 1986. "The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of personality and social psychology* (51:6), pp. 1173-1182.

Bliese, P. D. 2000. "Within-Group Agreement, Non-Independence, and Reliability: Implications for Data Aggregation and Analysis." pp. 349-381.

Borge Bravo, R., and Esteve Del Valle, M. 2017. "Opinion Leadership in Parliamentary Twitter Networks: A Matter of Layers of Interaction?," *Journal of Information Technology & Politics* (14:3), pp. 263-276.

Boutyline, A., and Willer, R. 2017. "The Social Structure of Political Echo Chambers: Variation in Ideological Homophily in Online Networks," *Political Psychology* (38), pp. 551-569.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. 2007. *Time Series Analysis: Forecasting and Control*. Hoboken, N. J.: John Wiley & Sons.

boyd, d., Golder, S., and Lotan, G. 2010. "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter," *43rd Hawaii International Conference on System Sciences (HICSS)*, Honolulu, HI, USA: IEEE, pp. 1-10.

Bradshaw, S., Howard, P. N., Kollanyi, B., and Neudert, L.-M. 2020. "Sourcing and Automation of Political News and Information over Social Media in the United States, 2016-2018," *Political Communication* (37:2), pp. 173-193.

Buchanan, T. 2020. "Why Do People Spread False Information Online? The Effects of Message and Viewer Characteristics on Self-Reported Likelihood of Sharing Social Media Disinformation," *Plos one* (15:10), p. e0239666.

Chen, J., Geyer, W., Dugan, C., Muller, M., and Guy, I. 2009. "Make New Friends, but Keep the Old: Recommending People on Social Networking Sites," *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 201-210.

Cheung, C. M., Wong, R. Y. M., and Chan, T. K. 2016. "Online Disinhibition: Conceptualization, Measurement, and Relation to Aggressive Behaviors," in: *International Conference on Information Systems (ICIS)*. Dublin, Ireland.

Coscia, M., and Rossi, L. 2020. "Distortions of Political Bias in Crowdsourced Misinformation Flagging," *Journal of the Royal Society Interface* (17:167), p. 20200020.

Del Vicario, M., Gaito, S., Quattrociocchi, W., Zignani, M., and Zollo, F. 2017. "News Consumption During the Italian Referendum: A Cross-Platform Analysis on Facebook and Twitter," *2017*

IEEE International Conference on Data Science and Advanced Analytics (DSAA): IEEE, pp. 648-657.

Dolan, P., Hallsworth, M., Halpern, D., King, D., Metcalfe, R., and Vlaev, I. 2012. "Influencing Behaviour: The Mindspace Way," *Journal of Economic Psychology* (33:1), pp. 264-277.

Drasch, B., Huber, J., Panz, S., and Probst, F. 2015. "Detecting Online Firestorms in Social Media," in: *International Conference on Information Systems*. Fort Worth, TX, USA.

Dubois, E., and Gaffney, D. 2014. "The Multiple Facets of Influence: Identifying Political Influentials and Opinion Leaders on Twitter," *American Behavioral Scientist* (58:10), pp. 1260-1277.

Dumitru, E.-A. 2020. "Testing Children and Adolescents' Ability to Identify Fake News: A Combined Design of Quasi-Experiment and Group Discussions," *Societies* (10:3), p. 71.

Egan, L. 2024. "What Kamala Harris' 'Call Her Daddy' Media Strategy Is Really About. Hint: It's Not About Mainstream Media's Egos." Retrieved November 8, 2024, from https://www.politico.com/news/2024/10/07/kamala-harris-media-strategy-call-her-daddy-00182781

French, A. M., Storey, V. C., and Wallace, L. 2023. "The Impact of Cognitive Biases on the Believability of Fake News," *European Journal of Information Systems*:forthcoming), pp. 1-22.

Gao, P., Lee, C., and Murphy, D. 2020. "Financing Dies in Darkness? The Impact of Newspaper Closures on Public Finance," *Journal of Financial Economics* (135:2), pp. 445-467.

George, J. F., Tilley, P., and Giordano, G. 2014. "Sender Credibility and Deception Detection," *Computers in Human Behavior* (35), pp. 1-11.

Gilovich, T., Keltner, D., and Nisbett, R. 2010. *Social Psychology*, (1 ed.). New York, NY: Norton & Company.

Golovchenko, Y., Buntain, C., Eady, G., Brown, M. A., and Tucker, J. A. 2020. "Cross-Platform State Propaganda: Russian Trolls on Twitter and Youtube During the 2016 Us Presidential Election," *The International Journal of Press/Politics* (25:3), pp. 357-389.

Graham, T. 2024. "Exploring a Post-Truth Referendum: Australia's Voice to Parliament and the Management of Attention on Social Media," *Media International Australia* (0:0), pp. 1-24.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., and Lazer, D. 2019a. "Fake News on Twitter During the 2016 Us Presidential Election," *Science* (363:6425), pp. 374-378.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., and Lazer, D. 2019b. "Fake News on Twitter During the 2016 Us Presidential Election," *Science* (363:6425), pp. 374-+.

Guess, A., Nagler, J., and Tucker, J. 2019. "Less Than You Think: Prevalence and Predictors of Fake News Dissemination on Facebook," *Science advances* (5:1), pp. 1-8.

Guo, L., A. Rohde, J., and Wu, H. D. 2020. "Who Is Responsible for Twitter's Echo Chamber Problem? Evidence from 2016 Us Election Networks," *Information, Communication & Society* (23:2), pp. 234-251.

Guo, L., and Zhang, Y. 2020. "Information Flow within and across Online Media Platforms: An Agenda-Setting Analysis of Rumor Diffusion on News Websites, Weibo, and Wechat in China," *Journalism Studies* (21:15), pp. 2176-2195.

Hafner, R., Elmes, D., and Read, D. 2019. "Exploring the Role of Messenger Effects and Feedback Frames in Promoting Uptake of Energy-Efficient Technologies," *Current Psychology* (38:6), pp. 1601-1612.

Hamm, M., and Spaaij, R. 2015. "Lone Wolf Terrorism in America: Using Knowledge of Radicalization Pathways to Forge Prevention Strategies," National Criminbal Justice Reference Service, pp. 1-28.

Heavey, C., Simsek, Z., Kyprianou, C., and Risius, M. 2020. "How Do Strategic Leaders Engage with Social Media? A Theoretical Framework for Research and Practice," *Strategic Management Journal*:41), pp. 1490–1527.

Heese, J., Pérez-Cavazos, G., and Peter, C. D. 2022. "When the Local Newspaper Leaves Town: The Effects of Local Newspaper Closures on Corporate Misconduct," *Journal of Financial Economics* (145:2), pp. 445-463.

Hox, J. J. 2010. *Multilevel Analysis: Techniques and Applications*, (2 ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Janze, C., and Risius, M. 2017. "Automatic Detection of Fake News on Social Media Platforms," *21st Pacific Asia Conference on Information Systems (PACIS)*, Langkawi, Malaysia, pp. 1-16.

Jöntgen, H. 2020. "Fueling the Firestorm: Effects of Social Capital on Users' Persuasiveness During Online Firestorms," *28th European Conference on Information Systems (ECIS)*, Virtual Conference, pp. 1-17.

Joslyn, M. R., and Haider-Markel, D. P. 2006. "Should We Really "Kill" the Messenger? Framing Physician-Assisted Suicide and the Role of Messengers," *Political Communication* (23:1), pp. 85-103.

Kassin, S. M. 1983. "Deposition Testimony and the Surrogate Witness: Evidence for a" Messenger Effect" in Persuasion," *Personality and Social Psychology Bulletin* (9:2), pp. 281-288.

Katz, E. 1957. "The Two-Step Flow of Communication: An up-to-Date Report on an Hypothesis," *Public opinion quarterly* (21:1), pp. 61-78.

Katz, E., and Lazarsfeld, P. F. 1966. *Personal Influence, the Part Played by People in the Flow of Mass Communications*. Transaction publishers.

Khalid, A., Moore, E., and Liasson, M. 2024. "The Npr Politics Podcast." *Call Her Daddy, The View and 60 Minutes: Kamala Harris wants to reach everyone* Retrieved October 8, 2024, from https://www.npr.org/2024/10/08/1210938239/podcast-call-her-daddy-kamala-harris-all-the-smoke-the-view

Kim, A., and Dennis, A. R. 2019. "Says Who? The Effects of Presentation Format and Source Rating on Fake News in Social Media," *MIS Quarterly* (43:3), pp. 1025-1039.

Kim, A., Moravec, P. L., and Dennis, A. R. 2019. "Combating Fake News on Social Media with Source Ratings: The Effects of User and Expert Reputation Ratings," *Journal of Management Information Systems* (36:3), pp. 931-968.

Kitchens, B., Johnson, S. L., and Gray, P. 2020. "Understanding Echo Chambers and Filter Bubbles: The Impact of Social Media on Diversification and Partisan Shifts in News Consumption," *MIS quarterly* (44:4).

Larson, K., and Watson, R. 2011. "The Value of Social Media: Toward Measuring Social Media Strategies," *32nd International Conference on Information Systems (ICIS)*, Shanghai, China, pp. 1-18.

Lee, J. 2015. "The Double-Edged Sword: The Effects of Journalists' Social Media Activities on Audience Perceptions of Journalists and Their News Products," *Journal of Computer-Mediated Communication* (20:3), pp. 312-329.

Lee, J. Y., and Sundar, S. S. 2013. "To Tweet or to Retweet? That Is the Question for Health Professionals on Twitter," *Health Communication* (28:5), pp. 509-524.

Leong, C., Pan, S. L., Bahri, S., and Fauzi, A. 2019. "Social Media Empowerment in Social Movements: Power Activation and Power Accrual in Digital Activism," *European Journal of Information Systems* (28:2), pp. 173-204.

Lewis, S. C., and Westlund, O. 2015. "Actors, Actants, Audiences, and Activities in Cross-Media News Work: A Matrix and a Research Agenda," *Digital journalism* (3:1), pp. 19-37.

Linvill, D. L., Boatwright, B. C., Grant, W. J., and Warren, P. L. 2019. ""The Russians Are Hacking My Brain!" Investigating Russia's Internet Research Agency Twitter Tactics During the 2016 United States Presidential Campaign," *Computers in Human Behavior* (99), pp. 292-300.

Linvill, D. L., and Warren, P. L. 2020. "Troll Factories: Manufacturing Specialized Disinformation on Twitter," *Political Communication* (37:4), pp. 447-467.

Lord, C. G., Ross, L., and Lepper, M. R. 1979. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence," *Journal of personality and social psychology* (37:11), pp. 2098–2109.

Lowry, P. B., D'Arcy, J., Hammer, B., and Moody, G. D. 2016. "'"Cargo Cult" Science in Traditional Organization and Information Systems Survey Research: A Case for Using Nontraditional Methods of Data Collection, Including Mechanical Turk and Online Panels," *The Journal of Strategic Information Systems* (25:3), pp. 232-240.

Luo, M., Hancock, J. T., and Markowitz, D. M. 2022. "Credibility Perceptions and Detection Accuracy of Fake News Headlines on Social Media: Effects of Truth-Bias and Endorsement Cues," *Communication Research* (49:2), pp. 171-195.

Maclean, J. C., Buckell, J., and Marti, J. 2019. "Information Source and Cigarettes: Experimental Evidence on the Messenger Effect," National Bureau of Economic Research.

Matherly, T., and Greenwood, B. 2024. "No News Is Bad News: The Internet, Corruption, and the Decline of the Fourth Estate," *MIS Quarterly* (48:2), pp. 699-714.

McCullagh, P., and Nelder, J. A. 1989. *Generalized Linear Models*, (2 ed.). London, UK: Chapman and Hall.

McFarland, L. A., and Ployhart, R. E. 2015. "Social Media: A Contextual Framework to Guide Research and Practice," *Journal of Applied Psychology* (100), pp. 1653-1677.

McQuail, D., and Windahl, S. 2015. *Communication Models for the Study of Mass Communications*. Routledge.

Metzger, M. J., Flanagin, A. J., and Medders, R. B. 2010. "Social and Heuristic Approaches to Credibility Evaluation Online," *Journal of communication* (60:3), pp. 413-439.

Moravec, P., Minas, R., and Dennis, A. R. 2019. "Fake News on Social Media: People Believe What They Want to Believe When It Makes No Sense at All," *MIS Quarterly* (43:4), pp. 1343-1360.

Muhammed, S. T., and Mathew, S. K. 2022. "The Disaster of Misinformation: A Review of Research in Social Media," *International Journal of Data Science and Analytics* (13:4), pp. 271-285.

Mutahi, P., and Kimari, B. 2020. "Fake News and the 2017 Kenyan Elections," *Communicatio*), pp. 1-19.

Mwachiro, M. 2024. "Here Are the Cable News Ratings for October 2024." Retrieved November 8, 2024, from https://www.adweek.com/tvnewser/cable-news-ratings-october-2024/

Neubaum, G., Rösner, L., Ganster, T., Hambach, K., and Krämer, N. C. 2018. "United in the Name of Justice: How Conformity Processes in Social Media May Influence Online Vigilantism," *Psychology of Popular Media Culture* (7:2), pp. 185–199.

Nguyen, C. T. 2020. "Echo Chambers and Epistemic Bubbles," *Episteme* (17:2), pp. 141-161

O'Hara, K., and Stevens, D. 2015. "Echo Chambers and Online Radicalism: Assessing the Internet's Complicity in Violent Extremism," *Policy & Internet* (7), pp. 401-422.

Oeldorf-Hirsch, A., and DeVoss, C. L. 2019. "Who Posted That Story? Processing Layered Sources in Facebook News Posts," *Journalism & Mass Communication Quarterly* (97:1), pp. 141-160.

Oeldorf-Hirsch, A., and DeVoss, C. L. 2020. "Who Posted That Story? Processing Layered Sources in Facebook News Posts," *Journalism & Mass Communication Quarterly* (97:1), pp. 141-160.

Pal, A., Chua, A. Y., and Goh, D. H.-L. 2019. "Debunking Rumors on Social Media: The Use of Denials," *Computers in Human Behavior* (96), pp. 110-122.

Pennycook, G., Cannon, T. D., and Rand, D. G. 2018. "Prior Exposure Increases Perceived Accuracy of Fake News," *Journal of experimental psychology: general* (147:12), p. 1865.

Pennycook, G., and Rand, D. G. 2019. "Fighting Misinformation on Social Media Using Crowdsourced Judgments of News Source Quality," *Proceedings of the National Academy of Sciences* (116:7), pp. 2521-2526.

PowerfulJRE. 2024. "Joe Rogan Experience." *Joe Rogan Experience #2219 - Donald Trump* Retrieved November 8, 2024, from https://www.youtube.com/watch?v=hBMoPUAeLnY

Reese, S. D. 2020. *The Crisis of the Institutional Press*. John Wiley & Sons.

Reynolds, W. M. 1982. "Development of Reliable and Valid Short Forms of the Marlowe-Crowne Social Desirability Scale," *Journal of clinical psychology* (38:1), pp. 119-125.

Riemer, K., and Peter, S. 2021. "Algorithmic Audiencing: Why We Need to Rethink Free Speech on Social Media," *Journal of Information Technology* (36:4), pp. 409-426.

Rimal, R. N., Lapinski, M. K., Cook, R. J., and Real, K. 2005. "Moving toward a Theory of Normative Influences: How Perceived Benefits and Similarity Moderate the Impact of Descriptive Norms on Behaviors," *Journal of health communication* (10:5), pp. 433-450.

Roth, P. L., Goldberg, C. B., and Thatcher, J. B. 2017. "The Role of Political Affiliation in Employment Decisions: A Model and Research Agenda," *Journal of Applied Psychology* (102:9), pp. 1286-1304.

Roth, P. L., Thatcher, J. B., Bobko, P., Matthews, K. D., Ellingson, J. E., and Goldberg, C. B. 2019. "Political Affiliation and Employment Screening Decisions: The Role of Similarity and Identification Processes," *Journal of Applied Psychology* (105:5), pp. 472-486.

Ruggiero, D. 2015. "The Effect of a Persuasive Social Impact Game on Affective Learning and Attitude," *Computers in Human Behavior* (45), pp. 213-221.

Ruijten, P. A. M. 2020. "The Similarity-Attraction Paradigm in Persuasive Technology: Effects of System and User Personality on Evaluations and Persuasiveness of an Interactive System," *Behaviour & Information Technology*), pp. 1-13.

Ryan, F., Impiombato, D., and Pai, H.-T. 2022. "Frontier Influencers: The New Face of China's Propaganda," Australian Strategic Policy Institute (ASPI), Canberra, Australia, pp. 1-56.

Sahelices-Pinto, C., and Rodríguez-Santos, C. 2014. "E-Wom and 2.0 Opinion Leaders," *Journal of Food Products Marketing* (20:3), pp. 244-261.

Salge, C., Karahanna, E., and Thatcher, J. B. 2022. "Algorithmic Processes of Social Alertness and Social Transmission: How Bots Disseminate Information on Twitter," *MIS Quarterly* (46), pp. 229-260.

Shao, C., Ciampaglia, G. L., Varol, O., Yang, K.-C., Flammini, A., and Menczer, F. 2018a. "The Spread of Low-Credibility Content by Social Bots," *Nature communications* (9:1), pp. 1-9.

Shao, C., Hui, P.-M., Cui, P., Jiang, X., and Peng, Y. 2018b. "Tracking and Characterizing the Competition of Fact Checking and Misinformation: Case Studies," *IEEE Access* (6), pp. 75327-75341.

Sherif, M. 1936. "The Psychology of Social Norms,").

Sherif, M. 1961. *Intergroup Conflict and Cooperation: The Robbers Cave Experiment*. Oklahoma, USA: University of Oklahoma. Institute of Group Relations.

Shin, J., Jian, L., Driscoll, K., and Bar, F. 2018. "The Diffusion of Misinformation on Social Media: Temporal Pattern, Message, and Source," *Computers in Human Behavior* (83), pp. 278-287.

Snijders, T. A., and Bosker, R. J. 2012. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, (2 ed.). Sage.

Sommariva, S., Vamos, C., Mantzarlis, A., Đào, L. U.-L., and Martinez Tyson, D. 2018. "Spreading the (Fake) News: Exploring Health Messages on Social Media and the Implications for Health Professionals Using a Case Study," *American journal of health education* (49:4), pp. 246-255.

Spohr, D. 2017. "Fake News and Ideological Polarization: Filter Bubbles and Selective Exposure on Social Media," *Business Information Review* (34), pp. 150-160.

Starbird, K., DiResta, R., and DeButts, M. 2023. "Influence and Improvisation: Participatory Disinformation During the 2020 Us Election," *Social Media+ Society* (9:2), p. 20563051231177943.

Sterrett, D., Malato, D., Benz, J., Kantor, L., Tompson, T., Rosenstiel, T., Sonderman, J., and Loker, K. 2019. "Who Shared It?: Deciding What News to Trust on Social Media," *Digital Journalism* (7:6), pp. 783-801.

Sunstein, C. R. 2018. *# Republic: Divided Democracy in the Age of Social Media*. Princeton University Press.

Tajfel, H. 1982. "Social Psychology of Intergroup Relations," *Annual review of psychology* (33:1), pp. 1-39.

Taylor, D. G. 1982. "Pluralistic Ignorance and the Spiral of Silence: A Formal Analysis," *Public opinion quarterly* (46:3), pp. 311-335.

Tepper, B. J., Moss, S. E., and Duffy, M. K. 2011. "Predictors of Abusive Supervision: Supervisor Perceptions of Deep-Level Dissimilarity, Relationship Conflict, and Subordinate Performance," *Academy of Management Journal* (54:2), pp. 279-294.

Toutenburg, H., and Shalabh. 2009. "Incomplete Block Designs," in *Statistical Analysis of Designed Experiments, Third Edition*. New York, NY: Springer New York, pp. 181-244.

Treem, J. W., and Leonardi, P. M. 2012. "Social Media Use in Organizations: Exploring the Affordances of Visibility, Editability, Persistence, and Association," *Communication yearbook* (36), pp. 143-189.

Tsai, W.-H. S., and Men, L. R. 2017. "Social Ceos: The Effects of Ceos' Communication Styles and Parasocial Interaction on Social Networking Sites," *New media & society* (19:11), pp. 1848-1867.

Vaast, E., Safadi, H., Lapointe, L., and Negoita, B. 2017. "Social Media Affordances for Connective Action: An Examination of Microblogging Use During the Gulf of Mexico Oil Spill," *MIS quarterly* (41:4), pp. 1179-1205.

Valdez, A. C., and Ziefle, M. 2018. "Believability of News - Understanding Users Perceptions of Fake News and the Effectiveness of Fact-Checking Badges," in: *Congress of the International Ergonomics Association (IAE)*. Florence, Italy: Springer, pp. 469-477.

van der Meer, T. G. L. A., Hameleers, M., and Ohme, J. 2023. "Can Fighting Misinformation Have a Negative Spillover Effect? How Warnings for the Threat of Misinformation Can Decrease General News Credibility," *Journalism Studies* (24:6), pp. 803-823.

Vargo, C. J., Guo, L., and Amazeen, M. A. 2018. "The Agenda-Setting Power of Fake News: A Big Data Analysis of the Online Media Landscape from 2014 to 2016," *New media & society* (20:5), pp. 2028-2049.

Vosoughi, S., Roy, D., and Aral, S. 2018. "The Spread of True and False News Online," *Science* (359), pp. 1146-1151.

Wade, J., Roth, P. L., Thatcher Jason, B., and Dinger, M. 2020. "Social Media and Selection: How Talking Guns, Doctors, and Mary Jane Influence Your Future," *MIS Quarterly* (44:3), pp. 1301-1357.

Walther, J. B., DeAndrea, D., Kim, J., and Anthony, J. C. 2010. "The Influence of Online Comments on Perceptions of Antimarijuana Public Service Announcements on Youtube," *Human communication research* (36:4), pp. 469-492.

Wang, M., Rao, M., and Sun, Z. 2020. "Typology, Etiology, and Fact-Checking: A Pathological Study of Top Fake News in China," *Journalism Practice*), pp. 1-19.

Wayne, S. J., and Ferris, G. R. 1990. "Influence Tactics, Affect, and Exchange Quality in Supervisor-Subordinate Interactions: A Laboratory Experiment and Field Study," *Journal of applied psychology* (75:5), p. 487.

Weeks, B. E., Ardèvol-Abreu, A., and Gil de Zúñiga, H. 2017. "Online Influence? Social Media Use, Opinion Leadership, and Political Persuasion," *International Journal of Public Opinion Research* (29:2), pp. 214-239.

West, D. M. 2017. "How to Combat Fake News and Disinformation." Retrieved November 11, 2028, from https://www.brookings.edu/articles/how-to-combat-fake-news-and-disinformation/

Westley, B. H., and MacLean Jr, M. S. 1957. "A Conceptual Model for Communications Research," *Journalism Quarterly* (34:1), pp. 31-38.

Wilson, E. J., and Sherrell, D. L. 1993. "Source Effects in Communication and Persuasion Research: A Meta-Analysis of Effect Size," *Journal of the academy of marketing science* (21:2), pp. 101–112.

Wong, R. Y., Cheung, C. M., and Xiao, B. 2018. "Does Gender Matter in Cyberbullying Perpetration? An Empirical Investigation," *Computers in Human Behavior* (79), pp. 247-257.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*, (2nd ed.). Cambridge, Massachusetts: MIT press.

Zimmer, F., and Reich, A. 2018. "What Is Truth? Fake News and Their Uncovering by the Audience," *5th European Conference on Social Media*, pp. 374-381.

Zollo, F. 2019. "Dealing with Digital Misinformation: A Polarised Context of Narratives and Tribes," *EFSA Journal* (17), p. e170720.

Appendix A

Politician (Rep. John Shane)

Account type

- *Avatar* provides the picture of the politician
- *Community* shows high follower numbers
- *About* identifies the page as governmental official individual

Perceived similarity (political)

- *Header* shows picture of Democratic or Republican symbolism (donkey/elephant)
- *Posts* about (personal) news or promotion of Democratic or Republican legislation

Additional information

- *Links* to other advocates

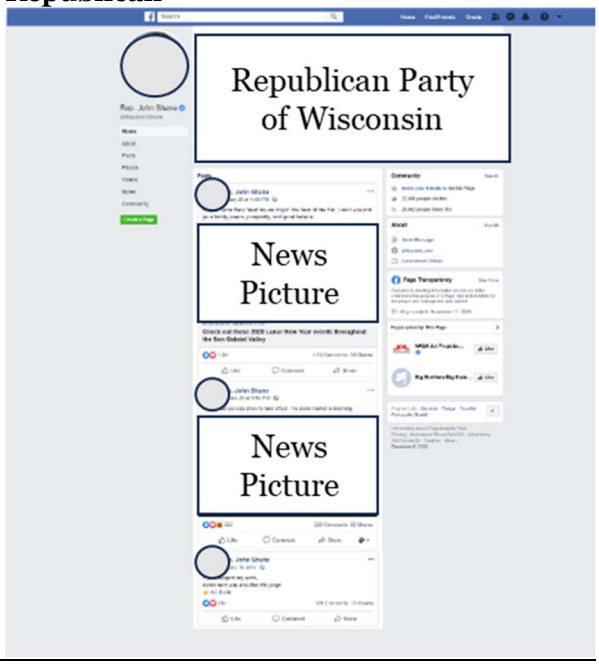
Democratic	Republican
	

Table 8. Democratic and Republican advocate manipulations

News Media (US National News (USNN))

Account type

- *Avatar* provides the logo of the news media
- *Community* shows high follower numbers
- *About* identifies the page as news media

Perceived similarity (political)

- *Header* provides a political figure with whom the news media politically identifies
- *News articles* with one article that provides talking points from either the Democratic (gun-regulations) or Republican (anti-abortion) perspective

Additional information

- *Links* to affiliated gatekeeper content (e.g., sports section of the new media)

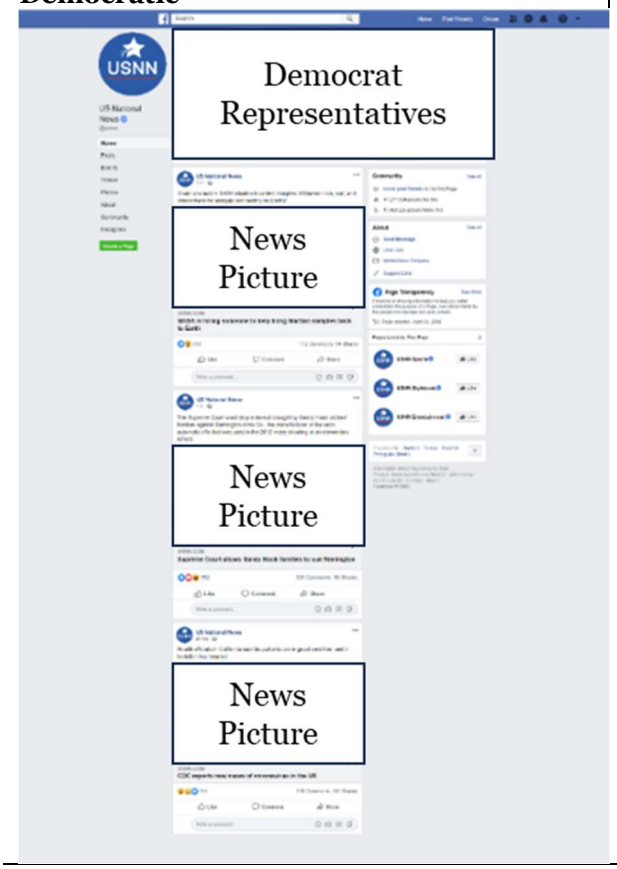
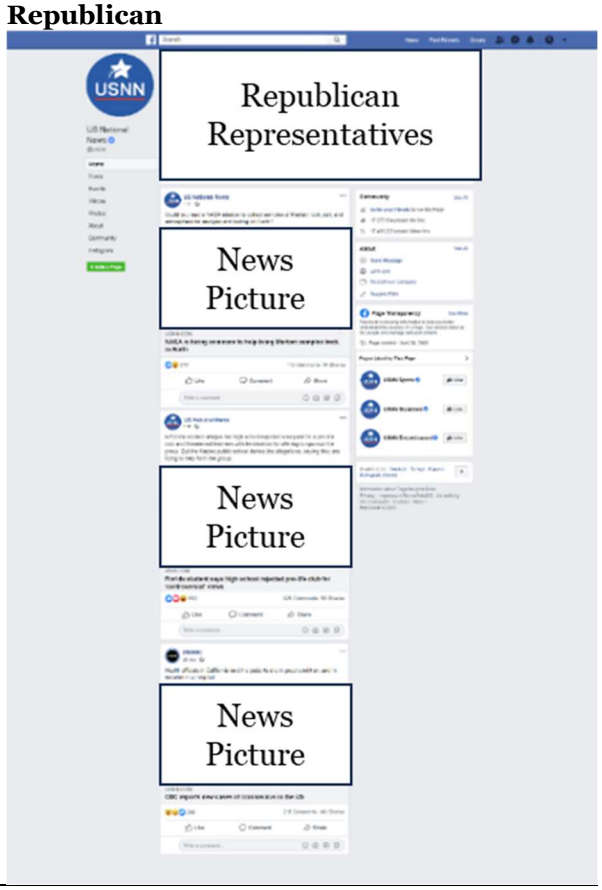
Democratic	Republican
	

Table 9. Democratic and Republican gatekeeper manipulations

User (Brian Smith)

Account type

- *Avatar* provides the picture of the user
- *Personal information* incorporates place of residence, job, education

Perceived similarity (political)

- *Header* shows picture of Democratic or Republican symbol (donkey/elephant)
- *Posts* about (personal) news or promotion of Democratic or Republican legislation

Additional information

- *Photos* without political affiliation
- *Friends*

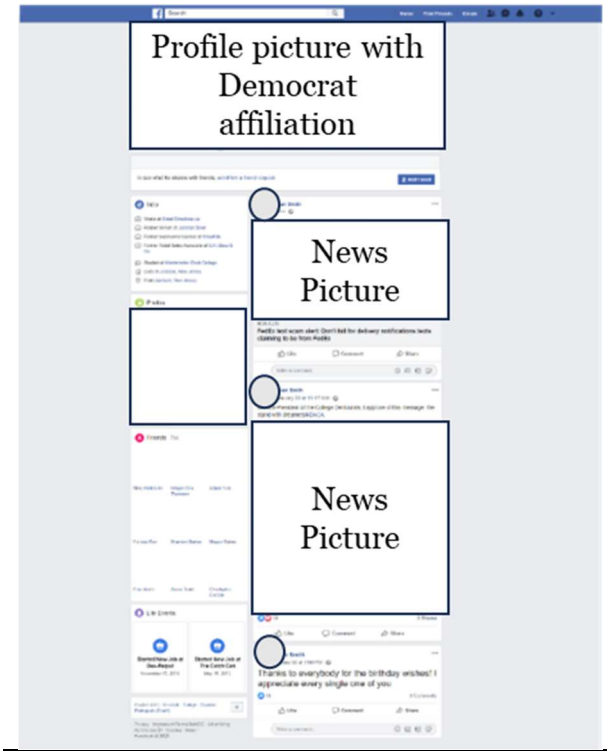
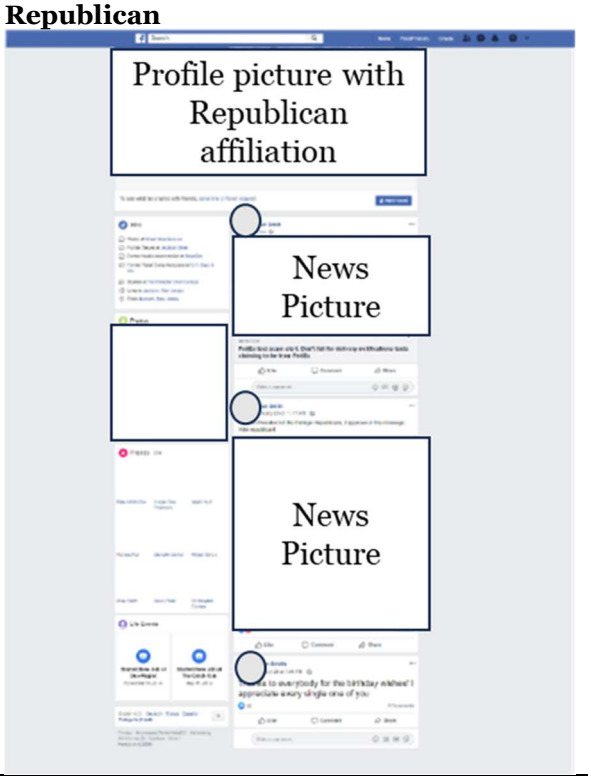
Democratic	**Republican**
	

Table 10. Democratic and Republican user manipulations

Appendix B

#	Fabricated News Story

NS1

NS2

NS 3

NS4

#	Fabricated News Story

NS5

NS6

NS7

NS8

#	Fabricated News Story
NS9	
NS10	
NS11	
NS12	

Table 11. Fabricated News Stories which were used in the study

Appendix C

Survey Measurements
Believability (Kim and Dennis 2019)
1. This post is truthful.
2. This post is believable.
3. This post is credible.
Perceived Similarity (Tepper et al. 2011) *(This person/news page and I…)*
1. are similar in terms of our outlook, perspective, and values.
2. assess issues in a similar way.
3. think alike in terms of proposing similar solutions for a problem.
4. think alike in a number of topics.
5. see things in much the same way.
Political Affiliation
1. Please tell us your political party affiliation.
Liking (Wayne and Ferris 1990)
1. How much do you like this person/news page?
Reputation (Kim and Dennis 2019)
1. [Stakeholder Name] is a trusted news source.
Topical Bias (Kim and Dennis 2019)
1. Do you find the issue described in the article important?
2. What is your personal position towards the article?
Social Desirability Bias (Reynolds 1982)
1. I never hesitate to go out of my way to help someone in trouble.
2. I have never intensely disliked anyone.
3. When I don't know something, I don't at all mind admitting it.
4. I would never think of letting someone else be punished for my wrongdoings.
5. I am always courteous, even to people who are disagreeable.
6. I sometimes feel resentful when I don't get my way.
7. There have been times when I felt like rebelling against people in authority even though I knew they were right.
8. I can remember "playing sick" to get out of something.
9. There have been times when I was quite jealous of the good fortune of others.
10. I am sometimes irritated by people who ask favors of me.

Table 12. Measurements as they were used in the online survey

Chapter II
Interaction

Paper IV
Beware of the Messenger!
How Social Media Account Characteristics Moderate the Confirmation Bias

Maximilian Haug
Neu-Ulm University of Applied Sciences

Marten Risius
Neu-Ulm University of Applied Sciences

Heiko Gewald
Neu-Ulm University of Applied Sciences

Jason Bennett Thatcher
University of Colorado Boulder

Unpublished

Beware of the Messenger!

How Social Media Account Characteristics Moderate the Confirmation Bias

Abstract

Social media allows its users to reach millions of other users instantly and without additional effort. This new form of mass communication enables misinformation campaigns to spread disinformation either as bot networks through coordinated inauthentic behavior or via the participatory disinformation of authentic users. To better understand how fake news campaigns leverage different types of accounts, we analyze the characteristics of these different accounts and evaluate how they affect other users' news believability. We conducted an experiment in which more than 900 participants rated the believability of 12 artificial news headlines posted on Facebook by different users (i.e., accounts). Those accounts hold different characteristics in terms of the type of user, communication role, and political affiliation. Thereafter, we analyzed how account characteristics moderate the influence of confirmation bias on believability. We found that the characteristics of the messenger (perceived similarity, politician account types) sharing a fake post exacerbated confirmation bias, rather than the characteristics of the creator of the original post. We show how messenger (i.e., the account sharing content from an originator) characteristics serve as mechanisms driving the dissemination of fake news on social media networks and call for more account-centric research on the strategies of fake news campaigns and their impact on the believability of news on social media.

Keywords:

news believability, messenger effect, coordinated inauthentic behavior, participatory fake news campaigns, social media accounts, confirmation bias

Introduction

Social media has transformed traditional communication, enabling individuals to reach millions of other users instantly and without additional effort (Jin et al. 2019). This transformation accelerates various phenomena, such as societal polarization, filter bubbles, and misinformation (Chan and Fu 2017; Lu et al. 2022; Srivastava et al. 2020). A well-documented case of misinformation happened in the 2016 U.S. presidential election when the Russian Internet Research Agency created almost 3,500 social media accounts to propagate their agenda. These fake accounts impersonated Republican (e.g., @March_For_Trump) and Democratic (e.g., @Blacktivists) users and produced around 8.6 million

tweets propagating false information (Linvill and Warren 2020) to increase the polarization between Republicans and Democrats (Lazer et al. 2018; Linvill et al. 2019).

While we know that malicious actors create fake accounts to share information, we know little about the role played by the specific characteristics of these accounts (i.e., different account types) in accelerating the spread of misinformation (Khan et al. 2021).

In this paper, we use the context of *fake news*, i.e., disinformation that is deliberately designed to mislead or misinform readers, to understand how different communication roles shape the believability of information (Vosoughi et al. 2018). *Communication roles* describe who generates and who disseminates information on social media (Vaast et al. 2017). Accounts that create original posts have the role of the *originator,* whereas accounts that forward, share, and distribute these original posts have the role of the *messenger* (Sterrett et al. 2019). The distinction between these roles is relevant as, we argue, social media users engage with messengers and originators based on how they perceive them as being similar. Consistent with the *similarity-attraction paradigm* (Ruijten 2020; Tajfel 1982; Wilson and Sherrell 1993), malicious actors assign specific characteristics to accounts to generate feelings of similarity to increase the believability in their fake news. For example, during the 2016 U.S. presidential campaign, the IRA capitalized on the similarity-attraction paradigm by setting up accounts that represented both Democratic (e.g., @JerStoner) and Republican (@johnlarsen) perspectives in order to polarize opinions through the distribution of fake news (Linvill and Warren 2020). By using accounts with a particular political affiliation, fake news campaigns can generate feelings of similarity and liking with party members as a mechanism to increase the believability of fake news.

Extant research segregates find different categories of actors creating and sharing information on social media. There are *opinion leaders* (OL) which are characterized by having a large number of followers and social support (Katz and Lazarsfeld 1964). On social media, OL typically have two expressions: OL who generate credibility from who they are (i.e., celebrities or political figures) and OL who generate credibility from what they do (i.e., news media organizations). Fake news campaigns utilize these two categories to spread fake news. Fake opinion leaders emulate individuals or organizations (e.g., politicians or political parties) who purposely communicate to modify the perceptions of a topic (Westley and MacLean Jr 1957). For example, the IRA created the @MARCH_FOR_TRUMP to emulate Trump's campaign organization. There are fake news media accounts like @TodayPittsburgh emulate reliable news organizations (Linvill and Warren 2020). Fake user accounts (e.g., social bots) emulate social media audience members (Linvill and Warren 2020) and amplify content shared by other social media users (Campan et al. 2017).

In this paper, we explore how social media account characteristics (politician vs. news media vs. user) shape the relationship from confirmation bias to news believability. While fake news and confirmation bias research has confirmed that similarity between previously held opinions and the content of fake news influences believability (Johnson and Kaye 2015; Kim and Dennis 2019; Moravec et al. 2018), less is known about the impact of the characteristics of the accounts communicating disinformation. This is important because as Khan et al. (2021) point out, fake news research needs to focus on the

environments in which different forms of information are embedded. Therefore, we ask: *How do the characteristics of social media accounts moderate the effect of confirmation bias on news believability?*

To answer this question, we executed an experimental study exposing 931 participants to three different types of mock Facebook accounts (politician, news media, user) of different perceived similarity (Republican, Democrat) that shared 12 artificial news headlines either as an originator (created original post) or as a messenger (shared an original post). We found evidence for a messenger effect, indicating that messenger characteristics exert a more powerful magnifying effect than originator characteristics on the relationship between confirmation bias and news believability. These findings shed light the on mechanisms through which account characteristics moderate the effects of confirmation bias on the spread of fake news. The remainder of this paper unfolds as follows: First, we review relevant work on fake news, confirmation bias, and account characteristics in relation to fake news campaigns. Then, we present the research model. Next, we offer a detailed explanation of the method, the context of the experiment, and the results. The research note closes with a discussion of research, limitations, and contributions.

Theoretical Background

Confirmation Bias Increases Believability

When assessing the believability of news, people often fall prey to confirmation bias. They believe in information that confirms their beliefs and ignore information that challenges them (Devine et al. 1990; Koriat et al. 1980). This effect is commonly aligned with social media users' desire to resolve *cognitive dissonance*, the mental unease or tension evoked by confronting challenging information (Festinger 1962). To resolve cognitive dissonance, users discount or challenge contradictory information and overemphasize reaffirming information (Simon 1979). Previous studies have connected confirmation bias to the dismissal of counterinformation from credible sources (Halbach et al. 2019; Moravec et al. 2018), refusing to flag information as false (Coscia and Rossi 2020), or completely discounting flags that are suggestive of fake news (Moravec et al. 2018). Moreover, information that aligns with readers' opinions has been found to increase confidence in views already believed to be true (Halbach et al. 2019).

While the similarity of opinions and confirmation bias have been extensively studied, we know little about whether or how the characteristics of the messenger account sharing news might exacerbate confirmation bias. This is important because evidence suggests that message characteristics, such as communicative role, similarity, and account type, shape the power of "participatory misinformation campaigns" (Starbird et al. 2019) and "coordinated inauthentic behavior" (Keller et al. 2020) to induce confirmation bias. In the following, we focus on messenger social media accounts and explore the role that their characteristics might play in amplifying confirmation bias.

Coordinated Inauthentic Behavior and Participatory Disinformation Campaigns

Successful fake news campaigns diffuse disinformation across social media networks (Starbird et al. 2019) by using botnets of fake accounts to *coordinate inauthentic behavior* or by using *participatory disinformation* networks of real politicians, news organizations, and regular users (Wilson and Starbird 2021). Whether accounts represent bots or people, we argue that in large online networks, the attributes of these accounts, such as their *communicative roles*, *perceived similarity, and account types,* affect the strength of the confirmation bias regarding news believability.

The communicative role (i.e., originator or messenger) of an account may shape its impact on news believability. The originator is an account that posts original information, while the messenger is an account that redistributes the information on social media through sharing features. George et al. (2014) showed that when an originator is credible, people struggle to identify disinformation. However, Sterrett et al. (2019) demonstrated that messengers are more important than originators for influencing trust in the content. This influence can be explained through the elaboration likelihood model (Petty and Cacioppo 1986); i.e., messengers are more proximal and therefore easier to process, in contrast to originators who are more distal. Consequently, it is important to consider the communicative role of accounts (i.e., as originator or messenger) in the news sharing chain to understand the potential impact of a given account on the believability of news.

Based on the similarity-attraction paradigm, cues embedded in account profiles or posts can shape *perceived similarity,* and thereby magnify news believability, because people are more attracted to people who are similar to them and are also more willing to engage with such individuals and more likely to believe the information, they share (Ruijten 2020; Tajfel 1982; Wilson and Sherrell 1993). Perceived similarity may strengthen the influence of confirmation bias in that people are more likely to readily seek out and interpret information presented by those whose beliefs they perceive to be similar to their own (Lord et al. 1979). In contrast, *perceived dissimilarity* can strengthen disconfirmation bias, making people more likely to be overly critical and suspicious of information presented by those whose beliefs they perceive to be different from their own (Edwards and Smith 1996; Wade et al. 2020). A commonly used mechanism to signal similarity is to include political affiliation in an account's information or posts (Pennycook and Rand 2019; Roth et al. 2019; Wade et al. 2020); in any case, social media accounts offer multiple possibilities for disinformation campaigns to manipulate actor characteristics and magnify confirmation bias across a large audience in ways that bypass journalistic integrity and codes of ethics that are typically taken for granted.

In disinformation campaigns, account types tend to map to one of three *communication types* (Linvill and Warren 2020; Starbird 2019; Starbird et al. 2019; Westley and MacLean Jr 1957). *Opinion leaders* in the traditional understanding are accounts that present themselves as people or organizations who are purposively advocating for a product, service, or ideology (e.g., influencers/advertisers, environmental activist or organization, political representative, or party), while using their social support to reach their audience (Katz and Lazarsfeld 1964; Marshall 2020; McQuail and Windahl 2015).

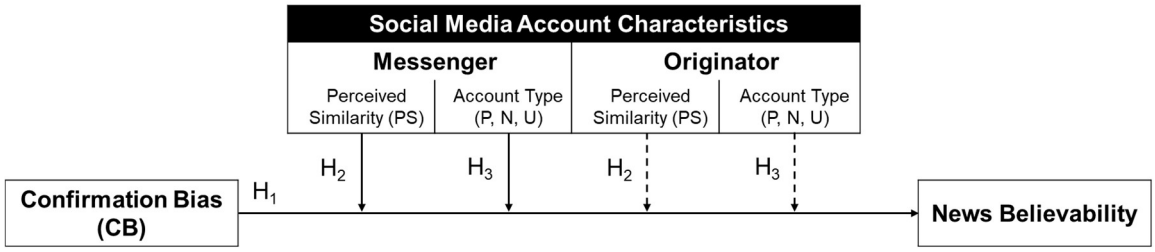
They are purposive in that they predominantly act in online networks to communicate and publicize their agendas in a particular direction (e.g., fashionable lifestyle, personal conviction, political position). For example, the Russian Internet Research Agency operated mock opinion leader accounts pretending to support a political view (e.g., Black Lives Matter, Make America Great Again) to appeal to a specific audience and polarize opinions (Linvill and Warren 2020). Opinion leaders clearly present their agenda upfront (e.g., account verifications, thematic profile text, avatar or picture, and so on) since they want to be recognized as sharing a consistent, purposeful message in online networks.

Within social media, news media have become also opinion leaders. These opinion leaders present themselves as credible nonpurposive news sources in online networks such as media organizations, reporters, news sites, editors, or experts (McQuail and Windahl 2015). On social media sites, malicious actors mimic known credible news organizations (e.g., pretending to be already defunct newspapers such as the *Chicago Daily News*). Malicious actors want these accounts to be perceived as nonpurposive accounts that share neutral, plausible information with interested audiences; thus, fake news media accounts select and transform information about events, which they then transmit to the target group via online networks (Westley and MacLean Jr 1957). To establish credibility, fake news media pretend to adhere to journalistic standards (e.g., veracity, accuracy, privacy) and communicate professionally (e.g., abstain from sensationalism).

User represents accounts that present themselves as authentic social media users seeking to acquire information, disseminate opinions, and participate in established networks of similar interests (Treem and Leonardi 2012). Users communicate with their friends and peers (e.g., chat) and within interest groups (e.g., group function on Facebook, subreddits), follow authentic accounts to personalize their feed (e.g., following), and disseminate information themselves (e.g., likes, shares). The audience influences future content in their feed mainly by engaging with content in the present (Bakshy et al. 2015). It constitutes the target audience for opinion leaders in online network communication.

Research Model

To understand how different roles and characteristics of messenger and originator accounts (i.e., perceived similarity and the type as politician, news media, or user) affect the relationship from confirmation bias to news believability, we assess how the interplay of account characteristics with confirmation bias impacts the believability of news.



Notes: P = Politician, N = News Media, U = User; dashed arrows indicate hypothesized non-significance

Figure 1. Research Model: How Social Media Account Characteristics Moderate Confirmation Bias

To evaluate the influence of account characteristics, we model confirmation bias as the main driver of believability (Halbach et al. 2019; Kim and Dennis 2019; Moravec et al. 2018). Based on the similarity-attraction paradigm and the related research findings on confirmation bias discussed above, we hypothesize:

H1: *Users are more likely to believe news that align with their preexisting beliefs about a topic.*

The effects of social media originators and messengers from the perspective of confirmation bias and the believability of news are yet to be explored. The elaboration likelihood model explains the circumstances under which stronger messenger or originator effects would be anticipated (Petty and Cacioppo 1986). Under a low elaboration likelihood, the characteristics of the more proximal messenger will be more impactful than those of the more distant originator. Hence, in the context of confirmation bias, we propose a messenger effect in which messenger characteristics (i.e., political party affiliation, account type) are more impactful for the believability of news than originator characteristics. Social media enables users to connect and engage with information from like-minded individuals (Bakshy et al. 2015). According to the similarity-attraction paradigm, knowledge of political affiliation would be expected to evoke feelings of similarity and shape the perceptions of the information that they share (Roth et al. 2017). Regarding social media, evidence suggests that disclosing political party affiliation indeed evokes feelings of similarity (Pennycook and Rand 2019), which, for example, have been found to bias recruiting decisions in favor of job candidates who are more similar to those empowered to make hiring decisions (Roth et al. 2019; Wade et al. 2020). Considering that, according to the elaboration likelihood model, messengers would generally be more proximal and originators more distal, we assume

that perceived similarity with the messenger is a stronger influence on the relationship between confirmation bias and believability than perceived similarity with the originator.

H2: *The moderating effect of perceived similarity on the relationship between confirmation bias and believability is stronger for messenger accounts than for originator accounts.*

To further differentiate between messengers and originators, we investigated the influence of account types. Previous research has suggested that accounts creating or disseminating fake news pose as politicians, news media, or users (Linvill et al. 2019; Linvill and Warren 2020). Consequently, we distinguish the opinion leaders (i.e., politicians and news media) from the users due to them having strong political positions and high followership and social status, compared to their audience members, and may use their status to reach and influence their target audience on social media (Borge Bravo and Esteve Del Valle 2017; Dubois and Gaffney 2014). Social media offers the possibility to pose as opinion leader. For example, fake news campaigns executed by Russian trolls leveraged the higher trustworthiness of the mentioned roles by creating fictional news media (e.g., Chicago Daily News) to make their news more credible (Linvill et al. 2019). We posit that the role of the messenger moderates the relationship between confirmation bias and believability more effectively than the role of the originator because, for messenger accounts, the primary attention is focused on account characteristics rather than on the characteristics of the original content. We thus hypothesize:

H3: *The moderating effect of opinion leaders on the relationship between confirmation bias and believability is stronger for messenger accounts than for originator accounts.*

Research Method

We conducted a 2x2x3 online experiment that presented made up news headlines from manipulated communicative role (messenger, originator), political similarity (similar, dissimilar), and account type (politician, news media, user) on a mock Facebook site.

Manipulations

Facebook allows for the creation of either “pages” (for organizations) or “profiles” (for individuals). We used Facebook “pages” for opinion leaders and “profiles” for individual users to create our manipulations. In our study, politician pages resembled a politician with a substantial follower base and proposed policy (Figure 2). News media pages resembled traditional news media and were also identified as a news organization by Facebook (Figure 3). News media were portrayed as an organization rather than an individual. User profiles as representative of the audience resembled private citizens and included information about hobbies and friends (Figure 3). We drew on related studies to manipulate political similarity (Roth et al. 2019; Wade et al. 2020); each account site included cues that highlighted Democratic or Republican Party affiliation. We used clear political symbolism, i.e., the donkey (for Democrats) and the elephant (for Republicans), as well as topics that are stereotypically associated with Democrats (e.g., pro-gun control) and Republicans (e.g., pro-life) (Figures 2-3).

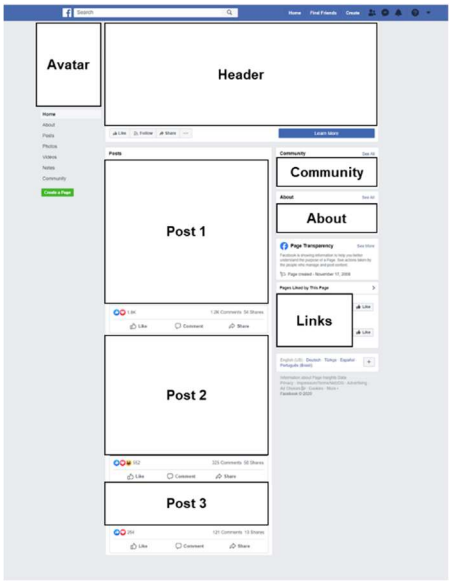
Facebook Page: Opinion Leader (politician and news media)	Manipulations
	**Account type** • *Avatar* provides the picture of the politician/logo of the news media • *Community* shows high follower numbers • *About* identifies the page as governmental official individual/news media **Perceived similarity (political)** • *Header* shows picture of Democratic or Republican symbolism (donkey/elephant)/ political figure with whom the opinion leader politically identifies • *Posts* about (personal) news or promotion of Democratic or Republican legislation/news articles with one article that provides talking points from either the Democratic (gun-regulations) or Republican (anti-abortion) perspective **Additional information** • *Links* to other opinion leader content (e.g., sports section of the news media)

Figure 2. Manipulations of Politician Acting as Republican or Democrat and News Media with Political Affiliation and Verified News Media

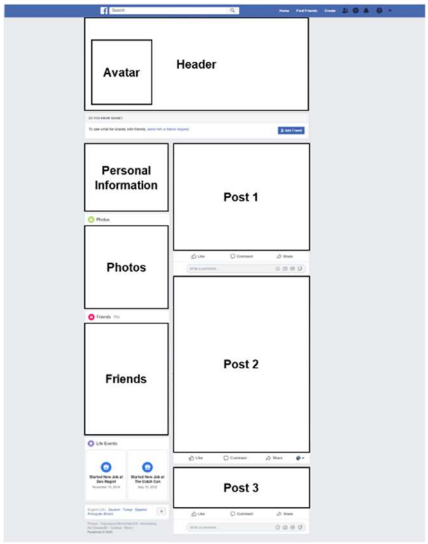
Facebook Profile: User	Manipulations
	**Account type** • *Avatar* provides the picture of the user • *Personal information* incorporates place of residence, job, education **Perceived similarity (political)** • *Header* shows picture of Democratic or Republican symbol (donkey/elephant) • *Posts* about (personal) news or promotion of Democratic or Republican legislation **Additional information** • *Photos* without political affiliation • *Friends*

Figure 3. Manipulations of User Acting as Pro-Republican or Pro-Democrat

Figure 4. Manipulation of the Communicative Role as Messenger or Originator

We established perceived (dis)similarity by exposing users to accounts that were equal to or different from their own political affiliation, and controlled for the communicative role by exposing the participants to different news article headlines that were posted by an originator and shared by a messenger (Figure 4). For the accounts, we reviewed real-world examples of the respective account types to establish ecological validity. We tailored the accounts accordingly by incorporating original content based on our review, and used generated names, avatars, and complementary information to ensure anonymity in this procedure. For further authenticity, information such as shares, likes, and nonrelated posts were added based on suggestions from the opinion leader literature (Borge Bravo and Esteve Del Valle 2017; Dubois and Gaffney 2014).

We created 12 news headlines based on examples from the literature (Moravec et al. 2018) (Table 1), choosing topics and articles that are not directly associated with either a Democratic or a Republican position that would be unlikely to induce a specific opinion or push the reader in a certain direction. This allowed us to focus solely on the similarity of accounts and the communication role.

#	News Headlines
1	ISIS Leader Calls for American Muslims to Support Women's March.
2	China to Discontinue all Trade with the United States.
3	Amazon in Contact with Airbus for Potential Amazon Airline.
4	U.K. about to Cut all Trade with the E.U while Focusing on the U.S. and Canada.
5	Animal Migration Doors to be Installed in Border Wall to Appease EPA.
6	United States Suing Volkswagen Over Cheating on Environmental Rules.
7	Alphabet (Google) about to get Sued for Record Sum for Mishandling Customer Data.
8	November Unemployment Rate Unexpectedly Rises to 6%.
9	Apple and Samsung Under Investigation Concerning Price Rigging in the Smartphone Market.
10	Daimler about to Shut Down Fabrication in the U.S.
11	Strategic Decision: Ford about to Step Back from Electric Mobility.
12	Microsoft about to Move Production out of China.

Table 1. Fabricated News Headlines, not Directly Associated with a Democratic or Republican Stance

Study Process

Our experiment followed a three-stage process—a pre-study, the main study, and a post-study. We followed a balanced incomplete block design to avoid reduplication of the same account types as messenger and originator (Toutenburg and Shalabh 2009). During the pre-study, we gathered the political affiliation of participants to ensure even distribution across the experimental conditions and an equal balance of Republican or Democratic participants. Participants who did not report affiliation with one of these two parties were disqualified. The participants were exposed to news headlines from originators and messengers in the main study. We introduced participants to the messenger the originator's profile pages. Participants rated each account's profile page according to similarity, liking, and reputation. To ensure that participants processed the information on the profile page and the political bias, they were asked an attention check question about the user's support of a topic found on their Facebook page or profile. Participants were then exposed to a sequence of six news posts from the originator that were shared by the messenger and were asked to rate their believability. Next, the participants were introduced to a different originator account and asked to rate it in the same fashion as before. The participants then proceeded to rate the believability of this originator's six news posts shared by the same messenger. We also used attention check questions while reading the news articles, such as checking a specific field during the first, the sixth, and the ninth news article. If the participant failed an attention check, their response was removed from the study. In the post-study, we asked questions about social desirability and debriefed participants.

The study took, on average, of 13 minutes to complete and was posted on surveymonkey.com. We recruited the sample on Amazon Mechanical Turk (mTurk), and each participant was compensated above minimum wage. Concerning mTurk, we followed the suggestions of Lowry et al. (2016). We recruited only mTurk participants that had earned the master's qualification, demonstrating high approval ratings and overall satisfactory task performance in the past. The participants were located in

the United States. Overall, we recruited 931 participants who rated 12 news headlines each (with six randomly missing data points due to nonresponses).

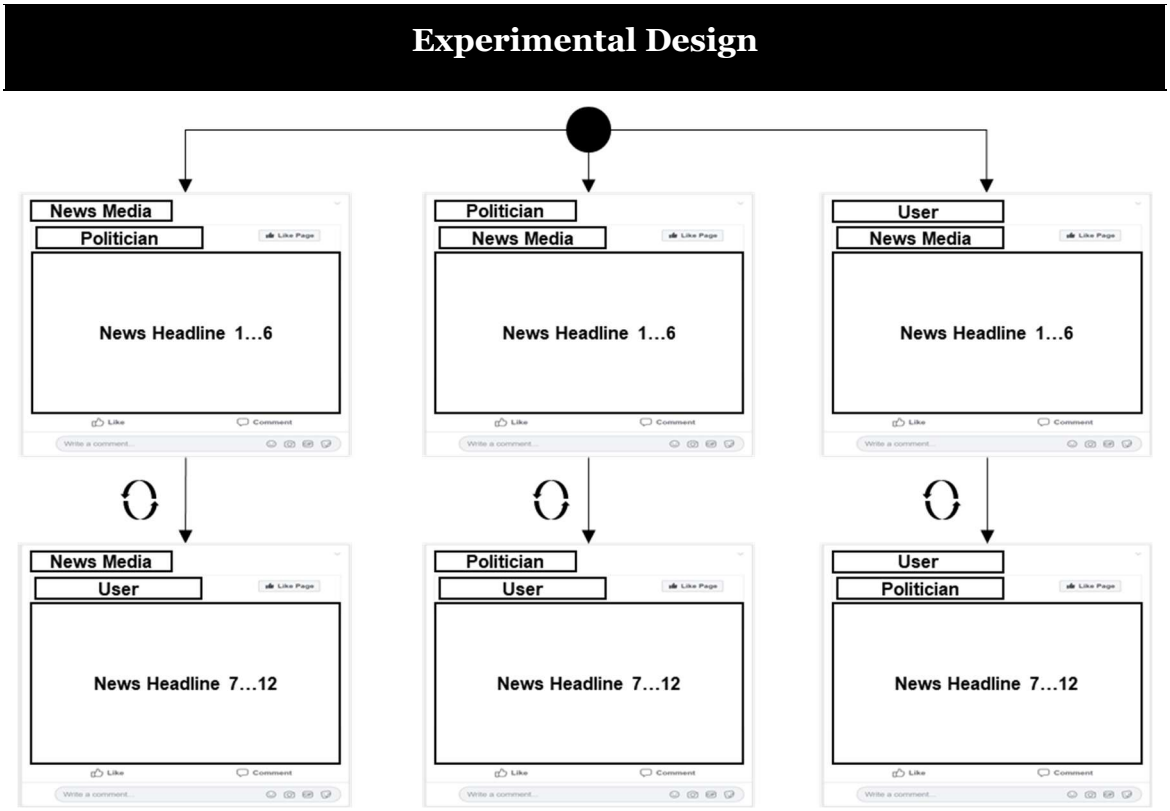


Figure 5. Study's Messenger-Originator Constellations of Varying Actor Types with Different Political Similarities

Operationalization and Measurements

We measured *believability* with three items on a seven-point scale ranging from "strongly disagree" to "strongly agree" in terms of whether the article was perceived as believable, credible, and truthful (Kim and Dennis 2019). For *perceived similarity*, we used a seven-point scale ranging from "strongly disagree" to "strongly agree," adapting Tepper et al. (2011) instrument to the context and the specific actor. We measured *liking* using the scale from Wayne and Ferris (1990) adapted to the context (person or Facebook page) with a five-point scale ranging from "1 = I don't like this person/site at all" to "5 = I like this person/site very much." Concerning *reputation*, following Kim and Dennis (2019), we asked participants whether or not they trusted the actor. Participants responded to "[Actor] is a trusted news source" on a seven-point scale ranging from "strongly disagree" to "strongly agree." Sterrett et al. (2019) found that both interest in the topic and familiarity with news have significant enough direct effects on trust and engagement to measure confirmation bias. Therefore, we measured the strength of the topical bias of the participants toward the article by multiplying the importance of the issue described in the article ("Do you find the issue described in the article important" 1 = not at all, 7 = extremely) with the personal position regarding the topic of the article ("What is your personal position towards the article?" -3 = extremely negative, +3 = extremely positive) in line with Kim and Dennis (2019). Lastly, we used

Reynolds (1982) established social desirability scale to establish whether a social desirability bias was present. The 10 items were measured on a seven-point scale ranging from "strongly disagree" to "strongly agree."

Analysis and Results

To test the hypotheses, we used nested multilevel random intercept, fixed-slope maximum likelihood regressions (Snijders and Bosker 2012). This enabled us to leverage the repeated measurements to account for the interblock error variance in the experiment's particular block design (Toutenburg and Shalabh 2009). Controlling for the multilevel structure also made our results independent from the individual participant's average level of believability (e.g., some participants are more likely to believe news than others) and the average believability of each news article (e.g., some news articles might be more believable than others), thus ensuring that the observed effects were related to the experimental manipulations of the account characteristics. Hence, we used a three-level model to refine our regression weight estimates by accounting for the nested structure of the repeated measurements of the individual instances of measuring news believability (Level 1), nested as repeated measurements in the individual participants (Level 2), which was nested through repeat measurements in the social media actor treatment groups (Level 3). In addition, we measured the participant's liking of the account and perceived reputation as control variables (Kim and Dennis 2019).

The intraclass correlation coefficients confirm that the proposed nested structure as the reliability of treatment-level means [ICC2] is expectably low at 0.1629. At the same time, the participant's variability is strongly dependent on their treatment group membership [ICC1] at 0.8943 (Bliese 2000). Additionally, we compared the mixed model with a standard OLS regression (Hox 2010). The results show a significant increase in model fit ($p < 0.001$) for the hierarchical structure compared to the standard OLS regression. Accordingly, we accepted the multilevel structure and evaluated the model by comparing their goodness-of-fit test statistics (McCullagh and Nelder 1989). Specifically, we computed the established restricted maximum log-likelihood values, Akaike's information criterion (AIC) (Box et al. 2007), and the Bayesian information criterion (BIC). Looking at the goodness-of-fit measures, the significantly decreasing model fit scores confirm the validity and appropriate complexity of our proposed model. Furthermore, R^2 = 23.58%, which indicates an acceptable amount of explained variance.

Political affiliation	Messenger			Originator		
	News Media	Politician	User	News Media	Politician	User
Similar	3.841 (.0334)	4.01 (.032)	3.848 (.035)	3.924 (.034)	3.858 (.034)	3.927 (.033)
Dissimilar	3.744 (.044)	3.797 (.04)	3.716 (.042)	3.761 (.041)	3.731 (.043)	3.757 (.042)

Note: n = 931 participants, standard deviations in parentheses below means

Table 2. Descriptive statistics for the treatment effects on news believability

The inspection of the descriptive statistics (Table 2) supports the assumed confirmation biasing effect because the average believability was higher for similar accounts ($\bar{x}_{Similar}$=3.901) compared to the dissimilar condition ($\bar{x}_{Dissimilar}$=3.751). These similarity-attraction effects do not appear to be clearly distinguishable on the general level of the communicative role, where originators' similarity changes are

associated with only slightly larger differences in news believability (0.153) compared to messengers (0.147). Instead, the effect of actor similarity on news believability seems to vary depending on account type in the following order: Believability changes appear to be strongest for politician (0.17), then for the audience user (0.151), and least for news media (0.13). Considering the interaction between communicative role and account type, these effects are particularly pronounced for messengers when (dis)similar politicians induce the biggest changes in believability (0.213), while the similarity of news media messengers have the weakest effect on believability (0.097).

To test the effects of the hypothesized individual fixed effects (Table 3), we calculated the corresponding Satterthwaite-corrected *t*-statistics and *p*-values for the path coefficients (Hox 2010). In addition, we performed omnibus tests in the form of simple contrasts of marginal linear predictions for multicategorical dummy variables or Wald tests for the interaction terms, respectively (Wooldridge 2010). The results support the hypothesized main effect of confirmation bias on the believability of news (b_{CB} = 0.05***, p < 0.001) (Hypothesis 1). We tested the messenger effect moderation hypotheses (Hypotheses 2 and 3) through the (cross-level) interaction terms with account types and perceived similarity (Baron and Kenny 1986). The pattern of results supports the prevalence of the messenger effect over that of the originator. As hypothesized, we found a significant moderator effect of perceived similarity with the messenger for the confirmation bias (b_{PSxCB} = 0.005, p < 0.001), but not with the originator (b_{PSxCB} = 0.000, p > 0.1) (Hypothesis 2). While the significant main effects of perceived similarity for both the originator (b_{CB} = 0.034, p < 0.05) and the messenger (b_{CB} = 0.036, p < 0.01) indicate the importance of perceived similarity, only the perceived similarity of the messenger exacerbates confirmation bias.

Results on the account type indicate a significant moderator effect of the account type (i.e., news media, politician, user) on believability, particularly for the messenger account type (Hypothesis 3). The originator account type does not have a significant main effect on news believability ($\chi^2_{(2)}$ = 0.32, p > 0.1), and no significant differences between different account types were found in the moderator condition. The significant omnibus Wald test for the originator ($\chi^2_{(5)}$ =73.52***) is related to the strong main effect of the confirmation bias without discernable moderating effects of the different account types. The messenger account type also has a highly significant moderator effect ($\chi^2_{(5)}$ = 83.88, p < .001), which—as the descriptive statistics suggest— is related to a tendentially significant difference between messenger politician and messenger news media ($b_{politician}$ = 0.01, p < 0.089), which we confirmed by rerunning the analysis with news media as the reference category. Even when the results deviated from the hypothesized effects in the sense that politicians showed stronger moderating effects while the believability of news media varied the least, these findings still confirm the assumed prevalence of a messenger effect because these differences appear in the messenger condition but not the originator condition.

Inspecting the predicted margins shows that politician and news media messengers have a slightly higher believability than user messengers, with a disordinal interaction between the opinion leaders. This means that compared to news media, politicians can exert stronger confirmation biasing effects when their fake articles support readers' existing opinions. However, readers are less likely to believe news articles that contradict their own views when they are shared by a politician versus news media.

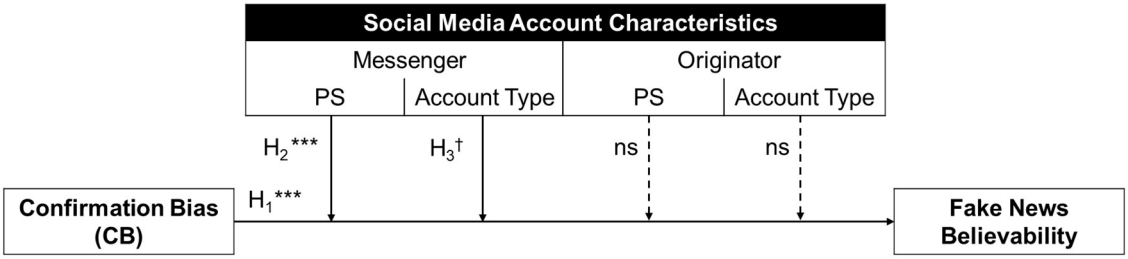
While we did not hypothesize differences between the two types of opinion leaders as messengers, the results support a moderating messenger effect that is pronounced for news media and politicians compared to users. We found that political affiliation has a statistically significant effect ($b_{Affiliation} = 0.05$, $p < 0.05$), indicating that Republicans are more likely to believe news. While this is in line with research that indicates differences between Democrats and Republicans and their engagement with news (Axt et al. 2020; Grinberg et al. 2019; Guess et al. 2019), we do not further elaborate, as it goes beyond the scope of this research note.

Fixed Effects	**Coefficient**		**Standard error**
Intercept	2.616***		.179
Controls			
Participant social desirability	.034**		.012
Participant political affiliation	.05*		.025
Originator liking	.027		.02
Originator reputation	.12***		.012
Originator perceived similarity (PS)	.034*		.013
Originator account type[a]		$\chi^2_{(2)} = 0.32$	
News Media	.001		.035
Politician	-.017		.035
Messenger liking	.063**		.02
Messenger reputation	.034**		0.011
Messenger perceived similarity (PS)	.036**		.013
Messenger account type[a]		$\chi^2_{(2)} = 1.95$	
News Media	.048		.036
Politician	.037		.035
Hypotheses			
H1: Confirmation bias (CB)	.053***		.008
H2: OriginatorPS x CB	.000		.001
H2: MessengerPS x CB	.005***		.001
H3: Originator Account Type x CB[b]		$\chi^2_{(5)} = 73.52^{***}$	
News Media	.003		.035
Politician	.008		.005
H3: Messenger Account Type x CB[b]		$\chi^2_{(5)} = 83.88^{***}$	
News Media	-.003		.005
Politician	.005		.005
Random part	Var. comp.		SE.
Level-three variance: News articles	.321		.131
Level-two variance: Participant	1.439		5.424
Level-one variance: Postings	.208		5.424
Log-likelihood (ML)[b]		-20,147.15***	
χ^2 (model improvement)[c]		$\chi^2_{(19)} = 1{,}971.26^{***}$	
AIC		40,302.31	
BIC		40,331.59	
R^2		.2358	

Note: Dummy coded variables (with reference category): political affiliation (Democrats), account type (news media); (CB) = confirmation bias, (PS) = perceived similarity, x = Interaction term, ML = maximum likelihood, AIC = Akaike information criterion, BIC = Bayesian information criterion.
Sample: News articles (level three) = 12; Participants (level two) = 931, Postings (level one) = 11,116.
p-values. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.09$ (two-tailed significance).
Parameter tests: [a]simple contrasts of marginal linear predictions: $\chi^2_{(df)}$, [b]Wald test: $\chi^2_{(df)}$,
[c]Likelihood ratio test:
-2ln(likelihood reference model) + 2ln(likelihood alternative model).

Table 3. Multilevel Regression Analysis Testing the Hypothesized Moderation Effects

Overall, the results indicate that distinguishing between the messenger and originator and their characteristics offers insight into news believability (Figure 6).



Note: PS = perceived similarity; p-values. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.09$ (two-tailed significance).

Figure 6. Research Model Results: Perceived Similarity and Account type of Messengers Moderate the Main Effect of the Confirmation Bias

Discussion

We took an account-centric approach to understand how communication roles, similarity, and account type influence the relationship from confirmation bias to news believability. We tested for a general messenger effect suggesting that accounts serving as salient messengers of a news headline are more impactful in terms of confirmation bias than the original news originator. We demonstrate that confirmation bias depends on the article's consistency with the reader's existing beliefs and similarity with the actor sharing the article. Our work reflects the relevance of perceived similarity for distinguishing among accounts that serve as messengers and originators of fake news in disinformation campaigns.

Our findings (Hypothesis 1) are consistent with prior research regarding the effects of confirmation bias on social media (Kim and Dennis 2019). We found support for communication roles shaping confirmation bias (Hypothesis 2)—that is, messengers increase the influence of confirmation bias on believability more strongly than news originators. We found that perceived similarity is important for the believability of news (for the originator and messenger) but that confirmation bias is exacerbated more by perceived similarity with the messenger. This interplay of the communication role and confirmation bias offers a better understanding of the drivers of the effectiveness of fake news campaigns (e.g., why only a few originator accounts are necessary to create fake news and many accounts are used to share it).

Hypotheses	Supported?
Confirmation bias	
H1: *Users are more likely to believe news that aligns with their preexisting beliefs about a topic.*	**Yes**
Messenger vs. originator characteristics	
H2: *The moderating effect of perceived similarity on the relationship between confirmation bias and believability is stronger for messenger accounts than for originator accounts.*	**Yes**
H3: *The moderating effect of opinion leaders on the relationship between confirmation bias and believability is stronger for messenger accounts than for originator accounts.*	**Partially supported**

Table 4. Hypotheses Results Overview

Our findings underscore the importance of differentiating between a similar messenger and originator, which extends findings from Sterrett et al. (2019), who were able to show that messengers have a strong influence on the trust that people place in the news. Therefore, the believability of news is a function of the article's content as well as how it is shared and who is sharing it.

The effect of the account type on the strength of confirmation bias was partially confirmed (Hypothesis 3). We found that the messenger account type can boost the effect of confirmation bias on the believability of news. This effect, however, is not equally strong for all opinion leaders as we hypothesized; it is particularly strong in the case of politicians. Politician messengers appear to be exceptionally effective in increasing the believability of news, especially when they are reaffirming previously held beliefs. When politicians contradict previously held opinions, people are less likely to believe them over news media (i.e., disconfirmation bias). This also means that fake politician messengers who share confirming or disconfirming views are particularly effective at polarizing discourse. This explains the efficacy of the Russian Internet Research Agency's strategy, which deployed mock account trolls that resembled Democratic and Republican Party affiliations to disinform and polarize the public discourse during the 2016 U.S. presidential election (Linvill et al. 2019). We believe that further research is needed to explore the impact of account types in this context. For example, future work, using a tool like qualitative comparative analysis, might be able to identify how the perceived similarity of configurations, account types (i.e., opinion leaders), and the communicative role of messenger (vs. originator) lead to confirmation bias when consuming news.

Theoretical Implications

This study has significant theoretical implications. We demonstrate the importance of distinguishing between the communication roles of accounts on social media. Although much fake news research currently seems to assume that the originator and messenger are the same (Kim and Dennis 2019; Kim et al. 2019), traditional communication research has established the importance of accounts used to source versus process information (Lee and Sundar 2013; Metzger et al. 2010; Wilson and Sherrell 1993). Since researchers have argued that opinion leaders do not form a structurally separate group from

the social systems of regular users (Borge Bravo and Esteve Del Valle 2017), our findings underscore the importance of considering communication roles to understand how to detect and prevent fake news dissemination (Vosoughi et al. 2018). Based on our findings, future research might consider how the interplay between account types (i.e., those of opinion leaders) and drivers of perceived similarity (e.g., political affiliation, lifestyle, or brand support) could impact the dissemination of content within an immediate community but also among interlinked online and offline communities.

Practical Implications

Our findings have important practical implications that suggest that combatting polarizing disinformation operations, like the one executed by the Russian IRA during the 2016 U.S. presidential election (Linvill et al. 2019) or the participatory disinformation campaigns during the 2020 U.S. presidential election (Wilson and Starbird 2021), requires focusing on messenger accounts. Given that it is difficult to predict what content will be posted, identifying and monitoring messenger accounts that redistribute fake news according to role, type, and affiliation could serve as an early warning system for detecting emerging fake news campaigns and could help platform owners limit disseminating polarizing messages by selectively using counter mechanisms (e.g., de-platforming, blocking, or flagging).

Limitations

Beyond complaints about external validity, complicated experiments such as ours are sometimes criticized for confounds that limit internal validity. To minimize the risk of carry-over effects, we randomized the sequences of fake news and accounts while maintaining the sequence of seeing the messenger first and then the originator. This allowed us to differentiate between roles, but participants may have paid more attention to the first profile (messenger) than the second (originator). Alternatively, participants may have become more familiar with the messenger because this role stayed the same across the 12 articles, whereas the originator changed after six articles. However, the descriptive statistics suggest that the opposite is true when comparing originators and messengers and the participant's comments do not indicate any such bias. Nonetheless, further research could test for these possibilities by replicating the study with a balanced complete block design of experimental account characteristic conditions.

Future Research

This study suggests a need to further investigate the interplay of communication roles and account types on the dissemination of fake news on social media and the interplay between originator and messenger and various sources of fake news (Sterrett et al. 2019). Future research should thus pursue a more nuanced understanding of account similarity and dissimilarity by, for example, exploring whether a similar messenger sharing news from a dissimilar originator could reduce the disconfirmation bias of the dissimilar originator.

Furthermore, to understand messenger effects, we need to better understand the extent to which fake news is shared. For example, how many messengers share the content of a single originator? Is it five or fifty or five hundred? And what effects do these numbers have on believability? This issue is complex because messengers may share news about the same issue but from different originators, as is common during global protests such as the Black Lives Matter movement. Understanding the opinion-forming effects of the sharing dynamics of social media and originator-messenger combinations is essential to fully understand how mass online events develop (Drasch et al. 2015; Jöntgen 2020).

Lastly, this study focuses on understanding how account characteristics moderate the relationship of confirmation bias and believability but leave unexamined the opinion-polarizing effect of fake news. The impact of opinion polarization on public discourse is considered a threat to modern society (Sunstein 2018). Polarization can lead to severe societal impacts such as terrorism or other violent events (Cheung et al. 2016; Hamm and Spaaij 2015; Wong et al. 2018). Future research could build on our findings to identify circumstances under which different accounts cause polarization and lead to harmful behaviors.

Conclusion

We evaluated whether social media account characteristics, particularly those of news messengers, exacerbate the effect of confirmation bias on believability. We found that account characteristics amplified confirmation bias, which increased when a messenger shared news headlines. Specifically, we found that the perceived similarity of a messenger can boost news believability. Our results also shed light on the interplay between news and the accounts that disseminate it (i.e., originator or messenger). Our work suggests the importance of realizing a more dynamic understanding of how accounts shape the believability of news and how perceptions of accounts lead to highly polarized opinions and divide society. Therefore, our work extends existing fake news research by showing that not only is the topic or content of a post relevant but the account itself plays an important role in the misinformation dissemination process.

References

Axt, J. R., Landau, M. J., and Kay, A. C. 2020. "The Psychological Appeal of Fake-News Attributions," *Psychological Science*), pp. 1-10.

Bakshy, E., Messing, S., and Adamic, L. 2015. "Exposure to Ideologically Diverse News and Opinion on Facebook," *Science* (348:6239), pp. 1130-1132.

Baron, R. M., and Kenny, D. A. 1986. "The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of personality and social psychology* (51:6), pp. 1173-1182.

Benkler, Y., Tilton, C., Etling, B., Roberts, H., Clark, J., Faris, R., Kaiser, J., and Schmitt, C. 2020. "Mail-in Voter Fraud: Anatomy of a Disinformation Campaign," Harvard University, pp. 1-49.

Bliese, P. D. 2000. "Within-Group Agreement, Non-Independence, and Reliability: Implications for Data Aggregation and Analysis." pp. 349-381.

Borge Bravo, R., and Esteve Del Valle, M. 2017. "Opinion Leadership in Parliamentary Twitter Networks: A Matter of Layers of Interaction?," *Journal of Information Technology & Politics* (14:3), pp. 263-276.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. 2007. *Time Series Analysis: Forecasting and Control*. Hoboken, N. J.: John Wiley & Sons.

Campan, A., Cuzzocrea, A., and Truta, T. M. 2017. "Fighting Fake News Spread in Online Social Networks: Actual Trends and Future Research Directions," *2017 IEEE International Conference on Big Data (Big Data)*: IEEE, pp. 4453-4457.

Chan, C.-h., and Fu, K.-w. 2017. "The Relationship between Cyberbalkanization and Opinion Polarization: Time-Series Analysis on Facebook Pages and Opinion Polls During the Hong Kong Occupy Movement and the Associated Debate on Political Reform," *Journal of Computer-Mediated Communication* (22:5), pp. 266-283.

Cheung, C. M., Wong, R. Y. M., and Chan, T. K. 2016. "Online Disinhibition: Conceptualization, Measurement, and Relation to Aggressive Behaviors," in: *International Conference on Information Systems (ICIS)*. Dublin, Ireland.

Coscia, M., and Rossi, L. 2020. "Distortions of Political Bias in Crowdsourced Misinformation Flagging," *Journal of the Royal Society Interface* (17:167), p. 20200020.

Devine, P. G., Hirt, E. R., and Gehrke, E. M. 1990. "Diagnostic and Confirmation Strategies in Trait Hypothesis Testing," *Journal of Personality and Social Psychology* (58:6), p. 952.

Drasch, B., Huber, J., Panz, S., and Probst, F. 2015. "Detecting Online Firestorms in Social Media," in: *International Conference on Information Systems*. Fort Worth, TX, USA.

Dubois, E., and Gaffney, D. 2014. "The Multiple Facets of Influence: Identifying Political Influentials and Opinion Leaders on Twitter," *American Behavioral Scientist* (58:10), pp. 1260-1277.

Edwards, K., and Smith, E. E. 1996. "A Disconfirmation Bias in the Evaluation of Arguments," *Journal of Personality and Social Psychology* (71:1), pp. 5-24.

Festinger, L. 1962. *A Theory of Cognitive Dissonance*. Stanford university press.

George, J. F., Tilley, P., and Giordano, G. 2014. "Sender Credibility and Deception Detection," *Computers in Human Behavior* (35), pp. 1-11.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., and Lazer, D. 2019. "Fake News on Twitter During the 2016 Us Presidential Election," *Science* (363:6425), pp. 374-378.

Guess, A., Nagler, J., and Tucker, J. 2019. "Less Than You Think: Prevalence and Predictors of Fake News Dissemination on Facebook," *Science advances* (5:1), p. eaau4586.

Halbach, P., Burbach, L., Nakayama, J., Plettenberg, N., Ziefle, M., Valdez, A. C., and Ieee. 2019. "Would I Lie to You? How Users Evaluate Faked Online Content Depending on Its Publication Type," in *2019 Ieee International Professional Communication Conference*. pp. 56-64.

Hamm, M., and Spaaij, R. 2015. "Lone Wolf Terrorism in America: Using Knowledge of Radicalization Pathways to Forge Prevention Strategies," National Criminbal Justice Reference Service, pp. 1-28.

Hox, J. J. 2010. *Multilevel Analysis: Techniques and Applications*, (2 ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Jin, S. V., Muqaddam, A., and Ryu, E. 2019. "Instafamous and Social Media Influencer Marketing," *Marketing Intelligence & Planning*).

Johnson, T. J., and Kaye, B. K. 2015. "Reasons to Believe: Influence of Credibility on Motivations for Using Social Networks," *Computers in human behavior* (50), pp. 544-555.

Jöntgen, H. 2020. "Fueling the Firestorm: Effects of Social Capital on Users' Persuasiveness During Online Firestorms," *28th European Conference on Information Systems (ECIS)*, Virtual Conference.

Katz, E., and Lazarsfeld, P. F. 1964. *Personal Influence, the Part Played by People in the Flow of Mass Communications*. Transaction publishers.

Keller, T., Graham, T., Angus, D., Bruns, A., Nijmeijer, R., Nielbo, K. L., Bechmann, A., Neudert, L.-M., Marchal, N., and Bradshaw, S. 2020. "'Coordinated Inauthentic Behaviour'and Other Online Influence Operations in Social Media Spaces," in: *AoIR Selected Papers of Internet Research*.

Khan, A., Brohman, K., and Addas, S. 2021. "The Anatomy of 'Fake News': Studying False Messages as Digital Objects," in: *Journal of Information Technology*.

Kim, A., and Dennis, A. R. 2019. "Says Who? The Effects of Presentation Format and Source Rating on Fake News in Social Media," *MIS Quarterly* (43:3), pp. 1025-1039.

Kim, A., Moravec, P. L., and Dennis, A. R. 2019. "Combating Fake News on Social Media with Source Ratings: The Effects of User and Expert Reputation Ratings," *Journal of Management Information Systems* (36:3), pp. 931-968.

Koriat, A., Lichtenstein, S., and Fischhoff, B. 1980. "Reasons for Confidence," *Journal of Experimental Psychology: Human learning and memory* (6:2), p. 107.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., and Rothschild, D. 2018. "The Science of Fake News," *Science* (359:6380), pp. 1094-1096.

Lee, F. L. 2018. "Social Media, Political Information Cycle, and the Evolution of News: The 2017 Chief Executive Election in Hong Kong," *Communication and the Public* (3:1), pp. 62-76.

Lee, J. Y., and Sundar, S. S. 2013. "To Tweet or to Retweet? That Is the Question for Health Professionals on Twitter," *Health Communication* (28:5), pp. 509-524.

Linvill, D. L., Boatwright, B. C., Grant, W. J., and Warren, P. L. 2019. ""The Russians Are Hacking My Brain!" Investigating Russia's Internet Research Agency Twitter Tactics During the 2016 United States Presidential Campaign," *Computers in Human Behavior* (99), pp. 292-300.

Linvill, D. L., and Warren, P. L. 2020. "Troll Factories: Manufacturing Specialized Disinformation on Twitter," *Political Communication* (37:4), pp. 447-467.

Lord, C. G., Ross, L., and Lepper, M. R. 1979. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence," *Journal of personality and social psychology* (37:11), p. 2098.

Lowry, P. B., D'Arcy, J., Hammer, B., and Moody, G. D. 2016. ""Cargo Cult" Science in Traditional Organization and Information Systems Survey Research: A Case for Using Nontraditional Methods of Data Collection, Including Mechanical Turk and Online Panels," *The Journal of Strategic Information Systems* (25:3), pp. 232-240.

Lu, Y., Wu, J., Tan, Y., and Chen, J. 2022. "Microblogging Replies and Opinion Polarization: A Natural Experiment," *MIS Quarterly* (46:4).

Marshall, P. D. 2020. "Celebrity, Politics, and New Media: An Essay on the Implications of Pandemic Fame and Persona," *International journal of politics, culture, and society* (33:1), pp. 89-104.

McCullagh, P., and Nelder, J. A. 1989. *Generalized Linear Models*, (2 ed.). London, UK: Chapman and Hall.

McQuail, D., and Windahl, S. 2015. *Communication Models for the Study of Mass Communications*. Routledge.

Metzger, M. J., Flanagin, A. J., and Medders, R. B. 2010. "Social and Heuristic Approaches to Credibility Evaluation Online," *Journal of communication* (60:3), pp. 413-439.

Moravec, P., Minas, R., and Dennis, A. R. 2018. "Fake News on Social Media: People Believe What They Want to Believe When It Makes No Sense at All," *MIS Quarterly* (43:4), pp. 1343-1360.

Pennycook, G., and Rand, D. G. 2019. "Fighting Misinformation on Social Media Using Crowdsourced Judgments of News Source Quality," *Proceedings of the National Academy of Sciences* (116:7), pp. 2521-2526.

Petty, R. E., and Cacioppo, J. T. 1986. "The Elaboration Likelihood Model of Persuasion," in *Communication and Persuasion*. Springer, pp. 1-24.

Reynolds, W. M. 1982. "Development of Reliable and Valid Short Forms of the Marlowe-Crowne Social Desirability Scale," *Journal of clinical psychology* (38:1), pp. 119-125.

Roth, P. L., Goldberg, C. B., and Thatcher, J. B. 2017. "The Role of Political Affiliation in Employment Decisions: A Model and Research Agenda," *Journal of Applied Psychology* (102:9), p. 1286.

Roth, P. L., Thatcher, J. B., Bobko, P., Matthews, K. D., Ellingson, J. E., and Goldberg, C. B. 2019. "Political Affiliation and Employment Screening Decisions: The Role of Similarity and Identification Processes," *Journal of Applied Psychology* (105:5).

Ruijten, P. A. M. 2020. "The Similarity-Attraction Paradigm in Persuasive Technology: Effects of System and User Personality on Evaluations and Persuasiveness of an Interactive System," *Behaviour & Information Technology* (40:8), pp. 734-746.

Simon, H. A. 1979. "Rational Decision Making in Business Organizations," *The American economic review* (69:4), pp. 493-513.

Snijders, T. A., and Bosker, R. J. 2012. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, (2 ed.). Sage.

Srivastava, K. C., Shrivastava, D., Chhabra, K. G., Naqvi, W., and Sahu, A. 2020. "Facade of Media and Social Media During Covid-19: A Review," *Int. J. Res. Pharm. Sci.*), pp. 142-149.

Starbird, K. 2019. "Disinformation's Spread: Bots, Trolls and All of Us," *Nature* (571:7766), p. 449.

Starbird, K., Arif, A., and Wilson, T. 2019. "Disinformation as Collaborative Work: Surfacing the Participatory Nature of Strategic Information Operations," *Proceedings of the ACM on Human-Computer Interaction* (3:CSCW), pp. 1-26.

Sterrett, D., Malato, D., Benz, J., Kantor, L., Tompson, T., Rosenstiel, T., Sonderman, J., and Loker, K. 2019. "Who Shared It?: Deciding What News to Trust on Social Media," *Digital Journalism* (7:6), pp. 783-801.

Sunstein, C. R. 2018. *# Republic: Divided Democracy in the Age of Social Media*. Princeton University Press.

Tajfel, H. 1982. "Social Psychology of Intergroup Relations," *Annual review of psychology* (33:1), pp. 1-39.

Tepper, B. J., Moss, S. E., and Duffy, M. K. 2011. "Predictors of Abusive Supervision: Supervisor Perceptions of Deep-Level Dissimilarity, Relationship Conflict, and Subordinate Performance," *Academy of Management Journal* (54:2), pp. 279-294.

Toutenburg, H., and Shalabh. 2009. "Incomplete Block Designs," in *Statistical Analysis of Designed Experiments, Third Edition*. New York, NY: Springer New York, pp. 181-244.

Treem, J. W., and Leonardi, P. M. 2012. "Social Media Use in Organizations: Exploring the Affordances of Visibility, Editability, Persistence, and Association," *Communication yearbook* (36), pp. 143-189.

Vaast, E., Safadi, H., Lapointe, L., and Negoita, B. 2017. "Social Media Affordances for Connective Action: An Examination of Microblogging Use During the Gulf of Mexico Oil Spill," *MIS Quarterly* (41:4).

Vosoughi, S., Roy, D., and Aral, S. 2018. "The Spread of True and False News Online," *Science* (359), pp. 1146-1151.

Wade, J., Roth, P. L., Thatcher Jason, B., and Dinger, M. 2020. "Social Media and Selection: How Talking Guns, Doctors, and Mary Jane Influence Your Future," in: *MIS quarterly Forthcoming*.

Wayne, S. J., and Ferris, G. R. 1990. "Influence Tactics, Affect, and Exchange Quality in Supervisor-Subordinate Interactions: A Laboratory Experiment and Field Study," *Journal of applied psychology* (75:5), p. 487.

Westley, B. H., and MacLean Jr, M. S. 1957. "A Conceptual Model for Communications Research," *Journalism Quarterly* (34:1), pp. 31-38.

Wilson, E. J., and Sherrell, D. L. 1993. "Source Effects in Communication and Persuasion Research: A Meta-Analysis of Effect Size," *Journal of the academy of marketing science* (21:2), p. 101.

Wilson, T., and Starbird, K. 2021. "Cross-Platform Information Operations: Mobilizing Narratives & Building Resilience through Both'big'&'Alt'tech," *Proceedings of the ACM on Human-Computer Interaction* (5:CSCW2), pp. 1-32.

Wong, R. Y., Cheung, C. M., and Xiao, B. 2018. "Does Gender Matter in Cyberbullying Perpetration? An Empirical Investigation," *Computers in Human Behavior* (79), pp. 247-257.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*, (2nd ed.). Cambridge, Massachusetts: MIT press.

Chapter II
Interaction

Paper V
Privacy is Important! Or not? – Commenting and Liking Under Confirmation Bias on Social Media

Maximilian Haug
Neu-Ulm University of Applied Sciences

Heiko Gewald
Neu-Ulm University of Applied Sciences

Proceedings of the *Americas Conference on Information Systems (AMCIS), 2023*,
Panama City, Panama
https://aisel.aisnet.org/amcis2023/vcc/vcc/11/

Chapter II
Interaction

Paper VI
Supporting Opinions to Fit in:
A Spiral of Silence-Theoretic Explanation for Establishing Echo Chambers and Filter Bubbles on Social Media

Maximilian Haug
Neu-Ulm University of Applied Sciences

Christian Maier
University of Bamberg

Heiko Gewald
Neu-Ulm University of Applied Sciences

Tim Weitzel
University of Bamberg

Internet Research, 2025, (35:7), pp. 30-51
https://www.emerald.com/insight/content/doi/10.1108/intr-03-2024-0413/full/html
DOI: 10.1108/INTR-03-2024-0413

Chapter II
Interaction

Paper VII
Social Media and Me:
How Community Identity Influences Click Speech

Maximilian Haug
Neu-Ulm University of Applied Sciences

Christian Maier
University of Bamberg

Heiko Gewald
Neu-Ulm University of Applied Sciences

Journal of Computer Information Systems, 2024, pp. 1-12
https://www.tandfonline.com/doi/full/10.1080/08874417.2024.2345106
DOI: 10.1080/08874417.2024.2345106

Chapter III
Ramifications

Paper VIII
Balancing Taking and Giving:
Contextualization of Social Support and Social Overload in Online Mental Health Communities

Maximilian Haug
Neu-Ulm University of Applied Sciences

Alisha Tikil
Neu-Ulm University of Applied Sciences

Heiko Gewald
Neu-Ulm University of Applied Sciences

Proceedings of the *European Conference on Information Systems* (ECIS), 2025, Amman, Jordan
https://aisel.aisnet.org/ecis2025/social_virtual/social_virtual/3

Chapter III
Ramifications

Paper IX
Content Creators on Instagram –
How Users Cope with Stress on Social Media

Maximilian Haug
Neu-Ulm University of Applied Sciences

Julia Reiter
Neu-Ulm University of Applied Sciences

Heiko Gewald
Neu-Ulm University of Applied Sciences

Telematics and Informatics Reports, 2024, (13), p. 100111
https://www.sciencedirect.com/science/article/pii/S2772503023000713
DOI: 10.1016/j.teler.2023.100111

Chapter III
Ramifications

Paper X
Anger and Sadness –
Coping Strategies to Manage Negative User Interaction on Instagram

Maximilian Haug
Neu-Ulm University of Applied Sciences

Julia Reiter
Neu-Ulm University of Applied Sciences

Heiko Gewald
Neu-Ulm University of Applied Sciences

Proceedings of the *European Conference on Information Systems* (ECIS), 2024, Paphos, Cyprus
https://aisel.aisnet.org/ecis2024/track15_social_ict/track15_social_ict/1/

Appendix

Publications

Scientific Journals (Peer Reviewed)

Haug, Maximilian and Maier, Christian and Gewald, Heiko and Weitzel, Tim (2025) Supporting opinions to fit in: a spiral of silence-theoretic explanation for establishing echo chambers and filter bubbles on social media. Internet Research, 35 (7). pp. 30-51. ISSN 1066-2243

Haug, Maximilian and Reiter, Julia and Gewald, Heiko (2024) Content creators on Instagram - How users cope with stress on social media. Telematics and Informatics Reports, 13. ISSN 2772-5030

Schroeder, Tanja and Haug, Maximilian and Georgiou, Andrew and Seaman, Karla and Gewald, Heiko (2024) Evidence of How Physicians and Their Patients Adopt mHealth Apps in Germany: Exploratory Qualitative Study. JMIR Mhealth Uhealth, 12. Digital Paper 48345. ISSN 2291-5222

Seerig, Kirsten H. and Haug, Maximilian and Maier, Alexander and Gewald, Heiko (2024) The Healing Power of Words: Examining the Effect of Communication Styles on Appreciation within the Hospital Setting. Procedia Computer Science, 231. pp. 305-310. ISSN 1877-0509

Haug, Maximilian and Maier, Christian and Gewald, Heiko (2024) Social Media and Me: How Community Identity Influences Click Speech. Journal of Computer Information Systems: JCIS, 64. ISSN 2380-2057

Schroeder, Tanja and Haug, Maximilian and Gewald, Heiko (2022) Data Privacy Concerns Using mHealth Apps and Smart Speakers: Comparative Interview Study Among Mature Adults. JMIR formative research, 6 (6). e28025. ISSN 2561-326X

Conferences and Workshops (Peer Reviewed)

Haug, Maximilian and Tikil, Alisha and Gewald, Heiko (2025) Balancing Taking and Giving: Contextualization of Social Support and Social Overload in Online Mental Health Communities. In (Proceedings of the) 33rd European Conference on Information Systems (ECIS), June, 12-18, 2025, Amman, Jordan

Haug, Maximilian and Finze, Nikola and Gewald, Heiko (2025) Unveiling the Digital Self – Exploring Avatar Identification in Online Gaming. In: (Proceedings of the) 13rd World Conference on Information Systems and Technologies (WorldCIST), April, 15-17, 2025, Florianopolis, Brazil.

Haug, Maximilian and Reiter, Julia and Gewald, Heiko (2024) Anger and Sadness — Coping Strategies to Manage Negative User Interaction on Instagram. In: (Proceedings of the) 32nd European

Conference on Information Systems (ECIS) “People First: Constructing Digital Futures Together”, June, 13-19, 2024, Paphos, Cyprus, Paper 1868. ISBN 9781958200100

Haug, Maximilian and Finze, Nikola and Gewald, Heiko and Salou, Tanja (2024) Exploring the Role of Emotions in Continuous Use of Activity Trackers. In: Conference on Health IT and Analytics (CHITA), May, 2-4, 2024, Washington D.C., USA.

Haug, Maximilian and Finze, Nikola and Gewald, Heiko and Salou, Tanja (2024) Happiness and Sadness - Exploring the Role of Emotions in Continuous Use of Activity Trackers. In: (Proceedings of the) Conference of the Italian Chapter of AIS (itAIS), October, 11-12, 2024, Piacenza, Italy.

Finze, Nikola and Trapp, Kateryna and Haug, Maximilian (2024) Losing Control! The Role of Sense of Agency in Using Highly Invasive AI Systems. In: (Proceedings of the) 45th International Conference on Information Systems (ICIS), December, 15-18, 2024, Bangkok, Thailand, Paper 1992.

Seerig, Kirsten H. and Haug, Maximilian and Maier, Alexander and Gewald, Heiko (2023) Informal Hierarchies - Ramifications for Nursing and Interprofessional Communication in German Hospitals. In: (Proceedings of the) 13th Conference on Health IT and Analytics (CHITA), May, 4-6, 2023, Washington D.C., USA.

Haug, Maximilian and Gewald, Heiko (2023) Privacy is Important! Or not? – Commenting and Liking Under Confirmation Bias on Social Media. In: (Proceedings of the) 29th Americas Conference on Information Systems (AMCIS) "Diving into Uncharted Waters" ; Vol. 11, August, 10-12, 2023, Panama City, Panama, Paper 1785.

Thaler, Fabian and Haug, Maximilian and Gewald, Heiko and Brune, Philipp (2022) The Context sets the Tone - A Literature Review on Emotion Recognition from Speech using AI. In: (Proceedings of the) 19th Conference of the Italian Chapter of AIS (itAIS), October, 14-15, 2022, Catanzaro, Italy.

Haug, Maximilian and Gewald, Heiko (2022) Privacy in the Pandemic - Commenting and Liking Behavior on Social Media Concerning COVID-19. In: (Proceedings of the) Conference on Health IT and Analytics (CHITA), March, 4-5, 2022, Washington, D.C., USA.

Haug, Maximilian and Schroeder, Tanja and Gewald, Heiko and Burth, Laura (2022) This Disease Scares Me! Influences on Adoption of Mobile Health Applications by Seniors. In: (Proceedings of the) International Conference on Information Technology & Systems (ICITS), February, 9-11, 2022, San Carlos, Costa Rica, pp. 140-149. (Lecture Notes in Networks and Systems; 414). ISBN 9783030962937

Schroeder, Tanja and Haug, Maximilian and Gewald, Heiko (2021) The Ambivalence of Mhealth Apps - A Qualitative Research on Techno-stressors of Mobile Health Applications. In: (Proceedings of the) 27th Americas Conference on Information Systems (AMCIS) "Digital Innovation and Entrepreneurship"; Paper 1543, August, 9-13,2021, Montreal, Canada (online).

Schroeder, Tanja and Schulz, Thomas and Haug, Maximilian and Gewald, Heiko (2021) A Conceptual Framework for a Digital Health Innovation Ecosystem. In: (Proceedings of the) 25th Pacific Asia Conference on Information Systems (PACIS); Paper 350, July, 12-14, 2021, Dubai, UAE.

Haug, Maximilian and Dahm, Maria and Gewald, Heiko and Georgiou, Andrew (2021) Just Talk to me - A Qualitative Study of Patient Satisfaction in Emergency Departments. In: (Proceedings of the) MedInfo'21 "One World, One Health - Global Partnership for Digital Innovation" ; Paper 306, October, 2-4, 2021, (virtual).

Haug, Maximilian and Lanza, Jaro and Gewald, Heiko (2021) Only If It Affects Me! The Influence of Privacy on Different Adoption Phases. In: (Proceedings of the) 42nd nternational Conference on Information Systems (ICIS) "Building Sustainability and Resilience with IS", December, 12-15, 2021, Austin, TX, USA, Paper 1935. ISBN 9781733632591

Schroeder, Tanja and Haug, Maximilian and Gewald, Heiko (2021) The difference between motivation and volition matters! - A qualitative study on mobile health application adoption. In: (Proceedings of the) 29th European Conference on Information Systems (ECIS) "Human values crisis in digitizing world", June 14-16, 2021, Marrakech, Marokko (online).

Haug, Maximilian and Gewald, Heiko and Brune, Philipp (2020) Retweeting in the Age of Fake News - A Cognitive Style Perspective. In: 17th Italian Chapter of the Association for Information Systems (itAIS) "Organizing in a digitized world: Diversity, Equality and Inclusion", October, 16 - 17, 2020, Pescara, Italy.

Haug, Maximilian and Rössler, Philipp and Gewald, Heiko (2020) How Users Perceive Privacy and Security Risks Concerning Smart Speakers. In: (Proceedings of the) 28th European Conference on Information Systems (ECIS), June, 15-17, 2020, Marrakech, Marokko (online). ISBN 9781733632515

Haug, Maximilian and Rössler, Philipp and Gewald, Heiko (2020) Identification and Influence of Perceived Risks on Smart Speaker Use Behavior. In: 15. Internationale Tagung Wirtschaftsinformatik (WI2020), 9.-11. März 2020, Potsdam, Germany, pp. 1325-1331.

Haug, Maximilian and Gewald, Heiko (2020) Feeling Younger? An Investigation of Cognitive Age on IT Use. In: (Proceeings of the) International Conference on Information Technology & Systems (ICITS), February, 5-7, 2020, Bogota, Columbia, pp. 320-329. (Advances in Intelligent Systems and Computing book series; 1137). ISBN 9783030406899

Schroeder, Tanja and Haug, Maximilian and Gewald, Heiko (2020) Technostress: a qualitative Research on Negative Health Impacts of Mobile Health Applications. In: (Proceedings of the) 11th Conference on Health IT and Analytics (CHITA), October, 8-10, 2020, Washington D.C., USA (online), pp. 20-21.

Haug, Maximilian and Wagner, Heinz-Theo and Gewald, Heiko and Burth, Laura (2019) Adoption of Mobile Health Applications against Alzheimer's Disease. In: 10th Conference on Health IT and Analytics (CHITA), November, 15-16, 2019, Washington D.C, USA.

Haug, Maximilian and Gewald, Heiko (2019) Why do I rate? - Shedding Light on the Factors Influencing the Participation on Physician Rating Websites. In: (Proceedings of the) 52nd Hawaii International Conference on System Sciences (HICSS), January 8 - 11, 2019, Grand Wailea, HI, USA. ISBN 9780998133126

Haug, Maximilian and Gewald, Heiko and Burth, Laura (2019) Why do I feel younger? – A critical view on cognitive age and IT adoption. In: The 5th International Workshop on Adults Use of Information and Communication Technologies in Healthcare (auICTH), November, 2019, Coimbra, Portugal, pp. 433-438. (Procedia Computer Science; 160). ISBN / ISSN 1877-0509

Haug, Maximilian (2019) Endorse the Source – The Impact of Information Assessment on News Sharing Behavior. In: (Proceedings of the) 25th Americas Conference on Information Systems (AMCIS), August 15-17, 2019, Cancun, Mexico. ISBN 9780996683180

Risius, Marten and Aydingül, Okan and Haug, Maximilian (2019) Towards an Understanding of Conspiracy Echo Chambers on Facebook. In: (Proceedings of the) 27th European Conference on Information Systems (ECIS), June, 8-14, 2019, Stockholm and Uppsala, Sweden. ISBN 9781733632508

Rockmann, Robert and Gewald, Heiko and Haug, Maximilian (2018) Equal Access for Everyone? A Digital Divide Cascade for Retired Senior Citizens. In: (Proceedings of the) 26th European Conference on Information Systems (ECIS), June, 23-28, 2018, Portsmouth, UK. ISBN 9781861376671

Haug, Maximilian and Gewald, Heiko (2018) Are you the Reason Fake News exists? Investigating News Sharing Attitude on Twitter. In: (Proceedings of the) 39th International Conference on Information Systems (ICIS), December, 13-16, 2018, San Francisco, USA. ISBN 9781510881044

Haug, Maximilian and Gewald, Heiko (2018) Friendly Equals Competent? - The Role of the Doctor-Patient Relationship in Physician Ratings. In: ACM SIGMIS Computers and People Research, Buffalo, USA.

Haug, Maximilian and Gewald, Heiko and Rockmann, Robert (2018) Grandma's New Tablet – The Role of Mobile Devices in Trying to Innovate in IT. In: The 9th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN-2018) / The 8th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare / Workshop 4th auICTH, November, 5-8, 2018, Leuven, Belgium, pp. 460-465. ISBN / ISSN 1877-0509 (Procedia Computer Science; 141)

Haug, Maximilian and Gewald, Heiko (2018) How to Rate a Physician? - A Framework for Physician Ratings and What They Mean. In: (Proceedings of the) 15th conference of the Italian Chapter of AIS (itAIS), October, 12-13, 2018, Pavia, Italy.

Haug, Maximilian and Gewald, Heiko (2018) Is 60 the new 40? An Assessment of the Measures of Cognitive Age vs. Chronological Age. In: 9th Conference on Health IT and Analytics (CHITA), October, 19-20, 2018, Washington, DC, USA.

Haug, Maximilian and Gewald, Heiko (2018) To eWOM, or not? Investigating the Restricting Factors to Provide Online Ratings for Physicians. In: (Proceedings of the) 24th Americas Conference on Information Systems (AMCIS), August, 16-18, 2018, New Orleans, LA, USA, pp. 3784-3793. ISBN 9781510870161

Haug, Maximilian and Gewald, Heiko (2017) What Satisfies the Patient? An Analysis of Factors Influencing Patients' Online Ratings. In: 8th Conference on Health IT and Analytics (CHITA), November, 3-4, 2017, Washington, DC, USA.