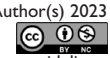
Misinformation rules!? Could “group rules” reduce misinformation in online personal messaging?

Article

new media & society
1–21



DOI: 10.1177/14614448231172964
journals.sagepub.com/home/nms

Misinformation rules!? Could "group rules" reduce misinformation in online personal messaging?

Andrew Chadwick, Natalie-Anne Hall and Cristian Vaccari
Loughborough University, UK

Abstract

Personal messaging platforms are hugely popular and often implicated in the spread of misinformation. We explore an unexamined practice on them: when users create "group rules" to prevent misinformation entering everyday interactions. Our data are a subset of in-depth interviews with 33 participants in a larger program of longitudinal qualitative fieldwork ($N = 102$) we conducted over 16 months. Participants could also donate examples of misinformation via our customized smartphone application. We find that some participants created group rules to mitigate what they saw as messaging's harmful affordances. In the context of personalized trust relationships, these affordances were perceived as making it likely that misinformation would harm social ties. Rules reduce the vulnerability and can stimulate metacommunication that, over time, fosters norms of collective reflection and epistemic vigilance, although the impact differs subtly according to group size and membership. Subject to further exploration, group rulemaking could reduce the spread of online misinformation.

Keywords

Affordances, correction, groups, metacommunication, misinformation, norms, rules, trust, vigilance

Corresponding author:
Andrew Chadwick, Online Civic Culture Centre, Department of Communication and Media, Loughborough University, Epinal Way, Loughborough LE11 3TT, Leicestershire, UK.
Email: a.chadwick@lboro.ac.uk

In this study, we focus on a previously unexamined practice that can affect the spread of misinformation online: the creation of "group rules" by people who want to keep misleading information out of their online personal messaging networks. We analyze a subset of in-depth interviews with 33 people who participated in a larger program of longitudinal qualitative fieldwork we conducted over 16 months (April 2021 to July 2022) with members of the UK public who broadly reflect the diversity of the UK population (Wave 1 $N = 102$, Wave 2 $N = 80$). Some participants also used a customized smartphone application to voluntarily send us examples of misinformation they saw on personal messaging. Where possible, we used this supplementary evidence to situate interview discussions in participants' everyday contexts of personal messaging use.

Our starting point is a conceptualization of messaging platforms as hybrid public-interpersonal communication environments. We theorize and explain how this distinctive context can encourage some people to create group rules. To briefly summarize our findings, we show how some people use rules to try to soften platform affordances they believe cause misinformation to spread, harm a group's members, provoke conflict, or derail a group's purpose. In smaller groups, some personal messaging users turn to rules because they believe social relationships in these contexts are driven by personalized trust springing from strong emotional bonds of kinship or friendship. Trustworthiness is perceived as inhering in the members of the group due to their close relationships and shared experiences; it is less dependent on whether information shared is externally verifiable. However, this social structure also makes misinformation more emotionally difficult to challenge, and as a result, social ties are more likely to be inadvertently exposed to harmful content than in other communication settings. Group rules are seen as one way to collaboratively reduce this vulnerability. In larger groups, such as those among work colleagues or neighborhood residents, there are also threshold barriers to speaking out; some participants perceive that group rules can help manage these. In larger groups, rules can serve to "institutionalize" the management of potentially harmful content. However, in these settings, lower levels of attentiveness and higher levels of delegation effects are also more likely to blunt the impact of rules.

Of course, rulemaking in online communities has a long history dating back to Internet's early days, when groups on Bulletin Board Systems, and later Usenet, developed ground rules such as FAQs (frequently asked questions) to regulate interactions among weak social ties.[1] To the best of our knowledge, however, this is the first examination of group rules for misinformation sharing on personal messaging. To be clear, 33 of the 102 participants in our broader project mentioned rules or rulemaking. Following an established approach in ethnography, we focus on a subsample of participants whose experiences we document in rich narrative detail in relation to this theme (Bird, 2003; Eliasoph, 1998; Hochschild, 2016; Pink and Mackley, 2013; Swart et al., 2019; Thorson and Wells, 2016; Toff and Nielsen, 2022). This is an exploratory study and our aim is to advance theory and reveal practices that could be examined further using other evidence and methods. Our conceptual framework differs from those informing most previous research into online misinformation. We situate the specificities of the messaging medium alongside affect, interpersonal and social relationships, and affordances and show how these forces converge to shape people's responses to misinformation. Our approach resonates with new scholarship that shows that these factors are increasingly

important in how citizens perceive news and public information (Masip et al., 2021; Tandoc et al., 2020; Toff and Nielsen, 2022). We extend these lines of prior research by developing new and distinctive theory and evidence related to personal messaging and how some people perceive that informal rules can protect against misinformation. More broadly, we respond to calls for new directions in research into how political conversations in everyday interpersonal interactions shape social attitudes and behavior (Eveland et al., 2011).

Online misinformation: personal messaging's distinctive role

WhatsApp (2020) has more than 2 billion users around the world. In the United Kingdom, more than 60% of the adult population use it regularly, which makes it more popular than any of the public social media platforms. Facebook Messenger has 18.2 million UK adult users (Office of Communications (Ofcom), 2021).

Over recent years, evidence has grown that personal messaging contributes to the spread of misinformation. Examples include financial scams, far-right conspiracy theories, sectarian religious myths, the promotion of fake remedies, and hate speech (Saurwein and Spencer-Smith, 2020). An emerging wave of research is grappling with how to study these platforms' impact on media systems, citizenship, and democratic norms (Kligler-Vilenchik, 2022; Malhotra and Pearce, 2022; Masip et al., 2021; Matassi et al., 2019; Pearce and Malhotra, 2022; Rossini et al., 2021; Valeriani and Vaccari, 2018). Yet how these services affect the spread of misinformation remains under-theorized and poorly understood (Pang and Woo, 2020). This is partly explained by the difficulties of gathering data. Messaging platforms mostly encrypt communication and lack public search archives. Arguably, this has led to neglect of messaging's unique patterns of use (Masip et al., 2021), which makes it all the more important to develop systematic, qualitative understanding of people's behavior on these platforms.

Research questions

Our study is exploratory and marries broadly inductive fieldwork with selected insights from prior research. Once we had discovered rulemaking was a practice among some of our participants, we set out two basic exploratory research questions to guide analysis. First, what leads some personal messaging users to perceive that spontaneous norms in habitual, everyday interactions on personal messaging are insufficient to protect their groups from misinformation? Second, when users try to create rules, how does the process work, and how do they perceive that the resulting rules affect the spread of misinformation?

Why this matters: everyday interactions, platform affordances, and social capacities

Contextual understanding of people's real-world relationships, their emotional responses to platform design, and the structure of group communication could generate new knowledge for reducing the spread of misinformation. This is especially relevant to personal

messaging platforms, which now mediate the production and reception of information for billions of people but are unsuited to the automated moderating and fact-checking typically used on *public* social media (e.g. Hameleers and Van der Meer, 2020). As we show, among our participants, rulemaking involved varying foci and degrees of formality, and its power differed depending on group size and membership, but the practice could have previously untapped advantages for combating misinformation. Our findings suggest that fact-checkers, news organizations, platform companies, and educators might further explore the benefits of encouraging group rulemaking in everyday personal messaging.

Personal messaging is important for everyday social and interpersonal relationships (Ofcom, 2021). A strand of audience research (e.g. Bird, 2003; Boczkowski et al., 2021; Pink and Mackley, 2013; Swart et al., 2019; Thorson and Wells, 2016) hints that these relationships layer into how people deal with misleading information and whether they feel emotionally empowered to speak out and challenge it.

Insight from human–computer interaction research on affordances also informs our framework (Masip et al., 2021). We use the concept of affordance in its phenomenological sense (e.g. Withagen et al., 2012). As Evans et al. (2017) write, an affordance "emerges in the mutuality between those using technologies, the material features of those technologies, and the situated nature of use" (p. 36). One way to tackle misinformation on personal messaging is to enhance people's awareness of how affect in social and interpersonal relationships converges with technological design to produce affordances that lead to misinformation's spread. Boosting people's capacities to develop group rules to counter these affordances can shape whether misinformation is challenged and corrected. This could scale up and have broad impacts.

Online personal messaging as hybrid public-interpersonal communication

Personal messaging can be understood as a hybrid public-interpersonal communication environment. This differs from how public social media blurs the boundaries between interpersonal and mass communication when interactions between people online are also simultaneously public for mass audiences to observe (Baym, 2015: 3). Services such as WhatsApp and Messenger are never fully public in that sense. WhatsApp's (2020) own data show that "90% of messages sent on WhatsApp are between two people, and the average group size is fewer than 10 people." These platforms are used mainly among strong-tie networks of family, friends, parents, co-workers, and local community members (Masip et al., 2021; Swart et al., 2019). Experiences are shaped by the iterative, mobile, and socially networked context of smartphone use and its affordance of perpetual, if sometimes ephemeral, everyday connection (Ling, 2012). Yet when (mis)information is shared, it has often originated in the more remote and public worlds of news, politics, and entertainment before it cascades across one-to-one and group settings. This can mean it loses markers of provenance (Bimber and Gil de Zúñiga, 2020), such as cues about its source, purpose, and temporality, along the way.

So, what makes personal messaging unique is not its privacy nor its intimacy. Both are important features of these platforms, but of greater significance for the spread of

misinformation is rapid and subtle switching—between private, interpersonal, and semi-public contexts and between one-to-ones, small groups, and larger groups. This happens in a context of relatively strong sender control over the audience for messages (Pearce and Malhotra, 2022). On personal messaging, information from the public world may intervene, but, in contrast with public social media, there are no fully public audience reception settings. An unbounded audience does potentially exist, afforded by the facility to join many groups or quickly "forward" messages into many other groups, but its size is hazy and impossible to predict because personal messaging use generates no meaningful searchability affordance.[2]

This hybrid public-interpersonal communication context affords (mis)information's easy transition from the public world into relatively private interpersonal communication networks, where different norms of correction might apply or where such norms may be absent. Rumors and misunderstandings in one-to-one or small-group interactions can also spread across other, larger messaging groups and acquire a more "public" character. And in these larger group contexts, weak norms of correction may apply, for example, due to social relationship factors such as people's anxiety about speaking out in larger groups, but also the technological ease of sharing and the difficulty of determining provenance.

From norms to rules and norms that valorize rules

For some people, group rulemaking can be a response to these vicissitudes of hybrid public-interpersonal communication. It offers a way to reduce the risks of misinformation spreading in their networks and causing harm, acrimony, and division. Rulemaking can also involve metacommunication (Coe et al., 2014; Molina and Jennings, 2018). Introducing rules to regulate the topics, information, or modalities of expression deemed permissible in a group can come to define the meaning and qualities of the relationships between its members. Rulemaking can become an end in itself: it valorizes collective reflection about rules. It can have beneficial effects by priming what Sperber et al. (2010) term "epistemic vigilance." By this, we mean metacommunication can heighten awareness of the need to evaluate plausibility and not rely only on the "goodwill" that Aristotelean theories of personalized trust argue is generated through shared experiences (McCroskey and Teven, 1999; see also Pasquetto et al., 2022).

Rules might also be used to route around how social and interpersonal relationships can make it difficult to challenge misinformation. Groups comprising family or close friends want to avoid continual conflict. This can lead to reticence toward challenging misleading messages, for fear of undermining the personalized trust that the group sees as inhering in friends and family members. On public social media, people tend to see as trustworthy the news and information they receive from family and friends (Dubois et al., 2020). As we show, with personal messaging, in some cases it was precisely *because* our participants recognized that family and friends trust each other implicitly that they perceived the need to develop ways to protect each other from misinformation. Rulemaking offers a way to protect the group by stimulating basic metacommunication without also undermining the goodwill ethic of care toward others that underlies personalized trust (McCroskey and Teven, 1999).

With larger groups comprising workmates or neighbors, social trust, which is more transactional (Uslaner, 2002), may be more important than personalized trust. However, in these contexts, there are still individual threshold barriers (Granovetter, 1978) to speaking out against misinformation. There is a perceived "cost" to crossing the threshold from inaction to action. Perceptions of these costs will vary between individuals but generally stem from a lack of confidence about marshaling evidence or not wanting to be perceived as undermining cohesion. In groups made up of larger proportions of people with weaker social relationships, rules reduce the need for ongoing confrontation. They serve to "institutionalize" or at least regularize the management of expression. However, in larger group settings characterized by lower cohesion, rules, though valuable in priming general vigilance, can lead to delegation effects: the rules come to be perceived as the responsibility of others to maintain and enforce. As we show, among our participants this could reduce attentiveness to the importance of the rules and therefore the rules' protective power.

Research design, data, and method

Our study sought to unearth people's experience and sense making—a goal difficult to achieve with digital trace data, surveys, or formal experiments. We conducted longitudinal in-depth interviews with members of the UK public whose characteristics reflected key features of the population (Wave 1 $N = 102$, Wave 2 $N = 80$, retention = 78.4%). In this study we analyze the experiences of the subset of 33 participants who told us about rule-making or rules during our interviews.

To recruit participants, we used an established agency, Opinium Research, which maintains a panel of more than 40,000 people who agree to participate in social and market research. To ensure balance in the sample, we deployed a short screening questionnaire (see the Supplementary Information file) and invited those who met the selection criteria (see section "Sampling" below) to participate in one-to-one interviews. We also invited participants to voluntarily upload relevant content encountered on personal messaging using a customized smartphone application. Interviewees granted informed consent and were compensated with a £35 payment for the first interview; £25 was provided for the second interview and any voluntary uploads. The latter occurred in the period between the first and second interviews and were self-selected examples of misinformation participants saw in their own personal messaging. Our aim was to use the uploads to enrich the second interviews. They allowed us to discuss examples that we could ask the participants to explain during the second interview, helping us situate their accounts within their everyday experiences (see the Supplementary Information file for the wording and further details). Loughborough University granted ethical approval before fieldwork began. The project design was finalized in 2019 and funded by a grant from the Leverhulme Trust in March 2020.

Sampling

Participants used at least one of the following at least a few times a week: WhatsApp, Facebook Messenger, iMessage, Android Messages, Snapchat, Telegram, Signal. As

Figure S1 (Supplementary Information) shows, WhatsApp and Facebook Messenger were clear leaders. Our recruitment screening ensured participants roughly reflected the UK population across gender, age, ethnicity, educational attainment, and basic digital literacy (Figure S2, Supplementary Information). This allowed us to balance in-depth, interpretive analysis with covering the experiences of a reasonably representative group of people from a variety of backgrounds and walks of life. One aim of the larger project (http://everyday-mis.info) was to examine neighborhood factors, so we recruited participants residing in three regions: London, the East Midlands, and the North East of England. To ensure balance on demographics and digital literacy, we employed an iterative strategy, with six recruitment and interview rounds on a rolling schedule for the first wave from April to November 2021. We then recontacted participants and invited them to second interviews, which ran from October 2021 to July 2022.

Procedure

Due to the Covid-19 pandemic, all interviews were conducted on Zoom, video-recorded, and fully transcribed. They were semi-structured, guided in part by themes we derived from pre-fieldwork pilot interviews. On average, each interview lasted about 1 hour 5 minutes. The aim was to capture attitudes, experiences, routines, how participants made sense of their social practices when using personal messaging, and how messaging layered into the texture of everyday life. We avoided *ex ante* definitions of misinformation, preferring participants to elucidate their own understandings. Where possible, we encouraged participants to check their smartphones during the interviews as an aide memoire. The 14-month duration of the fieldwork provided helpful variability of context and our method meant we had the ability to adapt as themes emerged organically, as was the case with rules or rulemaking. Longitudinal interviews enabled us to develop a richer understanding of how rulemaking originated and how it then evolved and shaped behavior.

Analysis

Using NVivo, we first conducted emergent and open interpretive coding of the transcripts (Corbin and Strauss, 1990). We were guided in part by our conceptual framework but remained open to themes that emerged. The project's principal investigator (first author) created the initial coding scheme. This was checked, discussed, and augmented by the other team members. Weekly team meetings were held to discuss consistency and add and refine codes over a period of about 12 weeks. We then moved to axial coding (Corbin and Strauss, 1990: 13) using NVivo's matrix coding query tool to explore intersections between themes. Finally, we moved to the selective coding phase by choosing rules as the central category and exploring how the 33 participants who discussed rules linked them with other themes (Corbin and Strauss, 1990: 14) and our conceptual framework. All team members agreed on the final coding scheme.

We report the social experience of some key participants with sufficient detail to enable readers to develop a rich, contextual understanding of how rulemaking played out in the experiences participants discussed with us. This ethnographic-narrative approach

fits with our focus on social relationships and everyday interactions (Eliasoph, 1998; Hochschild, 2016). Table S1 (Supplementary Information) reports selected further testimony from some participants not quoted in the next section. Interview material has been anonymized by removal or replacement of potential identifying details. All names are pseudonyms assigned by us.

Results

We begin with a discussion of one participant's experience, which encapsulates a key theme of our study: because group rules on personal messaging cannot derive from platform policies, they need to be developed through social interaction.

"There doesn't seem to be a sort of artificial intelligence, or even guidance to people who are running groups"

Ken is middle-aged and works in insurance in London. He recounted his experiences of setting up a neighborhood WhatsApp group for his apartment building, which included "22 flats." All residents were traditionally invited to an Annual General Meeting; some participated in the WhatsApp group. Ken became appalled by some of the residents' posts, which reflected various misperceptions about Covid and pandemic lockdown rules. This led to what he perceived to be unnecessary conflict over unfounded beliefs and misunderstanding of the government guidelines.

Ken intervened in the group to try to protect a resident who was medically vulnerable to Covid. He spoke of how, after this incident, he and the group then "made up rules as we went along." Ken is an avid user of public social media, and this experience taught him how they differ from WhatsApp, an environment facilitating communication in a liminal space between private and semi-public. Although the apartment building group was not especially large, with about 15–20 members, it was larger than most of the family and friend groups our participants mentioned. It was made up of some people with weaker ties than in a family or friendship group, but in other respects the ties were strong due to the fact that some participants shared apartments as romantic partners or families. Ken found it uncomfortable when he intervened and tried to enact what he described as "a bit of, sort of" moderation. This led him to make a revealing remark about how there was little to guide groups:

> *Ken:* I think it does make it more difficult: the bigger the group, the more difficult it is. Also, I think the group, you know, to avoid issues, the group almost needs a terms of reference, whether that's written or unwritten, that is very closely defined in terms of what, what it is trying to achieve. [. . .] Now, I guess, in that sense, I was doing a bit of sort of Facebook or WhatsApp moderation, which doesn't happen in real life, unless you happen to be called Donald Trump and then you'll get suspended, you know, by a committee [laughs]. But at a local level there doesn't seem to be a sort of artificial intelligence, or even guidance to people who are running groups as to what, what they should or shouldn't do.

Ken's remarks sprang from his rough-and-tumble experience of the apartment building WhatsApp group. In his second interview, he mentioned that differences of opinion in the group at times made him feel like he had "created a monster." He juxtaposed this somewhat plaintively with what he saw as the less difficult "artificial intelligence" (i.e. algorithmic sorting and automated moderation) he felt was routine on public platforms but absent from his group chat. At the same time, he recognized that was impossible to achieve on private messaging, so, in its absence, "even guidance" would help. The guidance in this case was the informal rules the apartment group developed after his suggestion. This experience was shared by Archie, a 66-year-old in the North East, who in his first interview spoke of how his local community WhatsApp group had similarly developed what he termed "blackballing" rules.

"We have this rule . . . absolutely no forwards!"

Some participants explained how they use rules to try to exert agency over what they and their friends and family saw as the potentially negative affordances of personal messaging platforms. A vivid example concerns a design feature much maligned in popular commentary for enabling misinformation to spread unchecked: the "forward."

Introduced by WhatsApp and Facebook Messenger to grow their services by encouraging people to easily repost material from one chat to multiple other chats, on personal messaging the forward is a well-known enabler of "viral" diffusion (Saurwein and Spencer-Smith, 2020). The problem is that it can be difficult to establish the origins of a forwarded post when it arrives into a chat. Unlike public posts on social media or, in fact, almost all other forms of digital communication, forwards on personal messaging present no identifying metadata to the user—not even the original poster's identity or the time of the original post. This becomes most problematic in the case of forwarded images and screenshots, but it also has implications for all kinds of shared information, including audio messages and hyperlinks.

A participant, Penelope, aged 20 and in the East Midlands, told us how her extended family WhatsApp group had decided on a "don't forward chain messages" rule aimed at protecting the grandparents and great grandparents in the group. According to her, this resulted in misinformation being shared "a lot less frequently." In the second interview, it became clear that this rule had empowered Penelope's parental generation—"mainly like my mum, my aunts, uncles"—to become active in "jumping forward whenever there are things that need to be corrected."

Another participant, Rehan, is in his late 20s, and one of his large family groups contains relatives in different countries, including the United Kingdom, Singapore, India, the United States, and Spain. Rehan recounted a recent episode when an uncle forwarded into the group a screenshot of an infographic detailing supposed infections and deaths from Covid-19. Rehan was concerned that the numbers in the screenshot were incorrect but he was unable to establish their provenance. He decided to post into the family group some links to official websites for pandemic statistics. His aim was to counter the impact of his uncle's "forward," which he was certain was misleading. Rehan then explained that the episode reminded him of why he and a subgroup of family members had decided

to create a separate family group with a simple foundational rule. The subgroup, with 32 members, was still large, but it had decided "absolutely no forwards! [. . .] we just say, particularly, no forwards in this group. Let's keep it just family content," he said.

Rehan explained that he and his family were aware that WhatsApp had introduced limits on how many times a post could be forwarded and a flag to highlight forwards "so you can sort of know if it's a viral thing or a mass thing." But, as he put it, he thought these restrictions "could be overcome easily [. . .] from a technical perspective" through copy and paste. For Rehan and his family, the chief concern was how the forward feature could interact with the context of trust and goodwill in the family group in ways that could render loved ones vulnerable to misleading information. As he described it, "forwards" created

> a lot of misinformation confusion—that people who think because their cousin sent it or their brother sent it or, you know, friends or somebody they have a relationship with sent it, it is true. Because obviously that's subjective, right, because you always want to side with people that you, you trust and care about, in your circle.

For Rehan and those in his family subgroup on WhatsApp, the "no forwards" rule was designed to mitigate how the affordances of easy sharing and loss of provenance could mislead others by preying on the relationships of trust in the group. In the family group context, trust was embodied in ties of kinship and prior relationships among the members, not derived from external cues of authority and verification. It did not need to be overtly maintained through references to external sources. Yet this was why Rehan felt the family group was vulnerable to misinformation. The "forward" feature made it more likely that a family member, trusted due to their social relationships in the group, perhaps time-pressed and wanting to warn or reassure, could too easily forward misinformation. Other group members would be more likely to accept it at face value as a result, so the "no forwards" rule was introduced.

"You shouldn't be sharing things like that . . . There needs to be a new group"

Luke is in his early 40s. In our first interview, he described how the main WhatsApp group in his workplace had recently undergone dramatic changes. By the time of our second interview, 7 months later, it was clear that the change had stuck and there was no going back to how the group was before the "new rules," as he described them.

Luke explained how the workplace WhatsApp group had become essential to the organization and coordination of work duties and opportunities in the firm. Alongside this, it had also been a lively forum for everyday discussion of news and events, a bit of workplace gossip, and what Luke said was "a lot of memes and GIFs and, like, humor and things," and "sort of banter." He described how two developments in the group's non-work chats had led to the decision to form a new group with new rules. First, during the early months of the pandemic, a growing number of colleagues were posting what Luke described as "Covid denier" misinformation. The second development came in the aftermath of the murder in May 2020 of George Floyd by police officer Derek Chauvin

in Minneapolis and the resulting resurgence of Black Lives Matter protests around the world. A co-worker had forwarded into the group a cartoon meme misrepresenting the Black Lives Matter movement, obliquely referring to Floyd's death and protests against historical statues of White slave owners. A few group members posted replies—"this is not the place to share this, you know, this isn't good"—Luke recalled. Accusations of racism followed, and the trade union launched an internal inquiry.

The outcome was that members of the union created a new main WhatsApp group. In their first post, they outlined the new group rules. As Luke explained,

> reps [. . .] sort of took over the running of the group and a new one was created with, like, rules that kind of have got to be followed and everything, going forward. So, you do, kind of, still get the occasional sort of sharing of things not work-related, but depending on what it, kind of, is, or if it's, could sort of cause offence and things, you do get people sort of being removed from the group. Like it's almost like the "naughty step" now, like they kind of get removed for a week.

No formal agreement was drawn up, but the first post by the new group's founders signaled that joining implied members were "agreeing to the rules" and they would remove colleagues who violated them. When asked in the interview about the new rules' focus, Luke elaborated, "keep sharing things to, like, work-related information primarily. And then I think it was like, you know, no kind of sharing of, like, GIFs or memes. And sort of being courteous to other people as well . . .."

Seemingly straightforward interactional and topic-based rules—be civil and only engage in work-related talk—also accreted a rule to mitigate the impact of the easy sharing and weak provenance affordances. Echoing Rehan's story of how the "no forwards" rule in his family group was aimed at misleading screenshot images, Luke's workplace group's "no GIFs or memes" rule stemmed from the experience of encountering visual misinformation and a misrepresenting and offensive meme in the group. The shift to rules was recognition of the potential power of visual representations and the ease with which they can travel across personal messaging networks, misinform, and cause conflict, division, and potentially harm.

It is also interesting to consider the power relations in Luke's story of the founding moment. It was not the result of bosses' decisions but an initiative by worker representatives to found a new group with new rules. The workplace WhatsApp group was already important for organizing tasks, and it was difficult for colleagues to exit it entirely. Still, this was not an official group run by managers. The act of founding a new group and announcing new rules was risky; colleagues might have left and established another group. But, as Luke explained later in the interview, not only did this process signal the importance of new sharing norms for the main workplace group, it also legitimized the founding process itself. This contributed to a shared ethic of responsibility after a period of acrimony caused by the misinformation.

When, in Luke's second interview, we returned to the topic, it quickly became clear that the new rules had made a difference. He described how "in terms of changes, I've kind of noticed that people are being a lot more cautious about what they're sharing. [. . .] People are just more concerned about it." Luke perceived that the new rules had

reduced misinformation in the group but also made him and his colleagues more aware of misinformation generally.

"Why are we arguing about this? It's just ridiculous"

Rehan's and Luke's experiences were echoed by Lydia. In this case, however, the group did not comprise family or workplace colleagues, but friends.

In her late 20s and working in an office job, Lydia described how her friendship messaging group had come to "an agreement we're just not gunna talk about Covid between us, unless it kind of particularly comes up." As she explained, this was "because it's just not, it's not a fun topic and everyone's sort of on slightly different wavelengths with it [. . .] and it was just creating a bit of tension."

The creation of new rules in Lydia's friendship group followed an argument about Covid lockdowns that had descended into back-and-forth sharing of conflicting information about infection rates and the dangers of the virus:

> that all kind of kicked off, and then, after that, uh, "debate," shall we call it, we sort of all said why are we like, why are we arguing about this? It's just ridiculous, like. There was a lot of, kind of, screenshots of the government website, everyone circling things and then sending it through and saying "oh what about this? what about this?"

This led to a "pretty outright agreement" to avoid discussing Covid. It meant the group as a whole avoided having to deal with the emotional and relational fallout from posts and screenshots some members deemed misleading or inaccurate.

Later, when the Covid vaccine became available, Lydia discovered that one of her friends was reluctant to take it. These views, and their supporting misinformation, were not spread in the group due to Covid having already been deemed off-limits—a positive outcome. On the negative side, however, the new rules meant clear opportunities to challenge Covid misinformation were lost. As was revealed in Lydia's second interview, her friend eventually got vaccinated, and although we have no direct evidence for this, we speculate that the "no Covid talk" rule could have reinforced to her friend the strong arguments against Covid misinformation that had led in the first place to the creation of the rule as a way to protect the friendships of the group's members. We return to the significance of this kind of trade-off in our concluding discussion.

"That wasn't the idea of the original group"

Sarah is in her early 50s and lives in east London. She participates in multiple messaging groups involving family and friends and belongs to a local neighborhood WhatsApp group. Sarah explained how the local group began when a neighbor put a note through people's front doors. It suggested they set up a support group to help with practical tasks such as shopping, childcare, DIY jobs, and advice about local services. The note through the door was this group's founding moment: it clearly referred to its purpose as a support network. Sarah, her partner, and several other neighbors signed up and began to post "practical information that we would share with one another." By the time of our second interview, the group had grown to "half the street," or "about 50" members. Some of these

people Sarah knew well and had "been in their homes." Others she knew well enough to greet on the street; others she would recognize when out walking, but not know by name.

Over time, Sarah and some other group members noticed posts drifting away from the group's original purpose. Some residents started to share news—some of it accurate, some misleading. In her eyes, the group was becoming less about help and support and more about "spreading fear, even, like, true or not [. . .] there were a few people that would post things that kind of raised your heart rate, made you kind of, oh my gosh!" As she saw it, "That wasn't the idea of the original group." The group had its topic-based rules, but they were being lost as it morphed into something less supportive that stimulated negative emotions. Sarah's account of what happened next was revealing.

Her first instinct had been to tell the group members "don't share that here." She explained, however, that the neighbor who started the group eventually intervened to remind members of the rules, but did so while avoiding overt confrontation:

> the woman that started it had a very good way of saying, you know, "I understand, we all understand, you're fearful. Maybe that could be shared in a different form." But, you know, she was very tactful [. . .] and actually I commended her for [it] about a week later. I ran into her, and I said, you know, you did, well done, for the way you handled that, kind of acknowledging that they were having a fear, but it wasn't quite the right place to share that, without kind of shutting them down and being rude, which is what I wanted to do, but I didn't do anything [. . .] Everyone else seemed to fall in line, [. . .] everyone else just kind of organic, organically started to understand what it was for and what it wasn't for.

The founder's intervention brought the local support group back to the rules and prevented it becoming populated with fear-inducing misinformation and conflict. It had a long-term impact as well. In Sarah's revealing language, it led to what she called an "organic" understanding of the group's purpose. In common with Luke's experience discussed earlier, it also stood the test of time. Nine months later, in her second interview, Sarah relayed how the neighbor who originally set up the group with the notes through doors had moved home and no longer lived nearby. And yet, Sarah said, the group was still lively and its rules had prevailed. They had

> just kind of taken, it's just sort of understood [. . .] so it's just kind of taken a natural, it's got a natural kind of basis, or so [chuckles] you know. So there doesn't really need to be anyone monitoring it.

This experience was shared by Barry, a Londoner in his early 40s, who in his first interview told us he belonged to a local school parents' WhatsApp group. In the second interview he discussed how the two parents who set up the group had started to say "this is not the forum" for off-topic posts, and this had a long-term impact on the group: "people tend to behave themselves," said Barry.

The fragility of rules in larger groups

Earlier we briefly noted that Archie, a participant in the North East, told us about his local community WhatsApp group's blackballing rules. As Archie expanded on his

experiences, it emerged that he had a rather hazy understanding of how those rules were enforced. In his first interview, he said,

> there seems to be a moderator [. . .] I'm assuming there's a moderator there, but I don't know that for a fact, you know. Maybe the person that set it up has some control, I don't know [. . .] it's just. . . that's the rule and that's what happens,

In the second interview, Archie explained that, in the time since his first interview, the local WhatsApp groups had in fact moved to become groups on Facebook, where he said moderation and information sharing (about local town planning issues) were clearer and "easier." This raises the issue of how group size and membership could matter for the power of rules. We now examine this in our final example, which concerns a large community-based group in a large city. Here, the underlying social relationships were diverse and complex, thus complicating the functioning of group rules. Some members knew each other from their interactions in the group, but overall the ties were relatively weak, and there were posts from people who were not well known to the group.

"I need to be more careful"

Priya is in her early 30s and works in a part-time IT job. She belongs to a community WhatsApp support group of people with links to the region outside the United Kingdom from where her family originates. The group is an offshoot of a now-defunct group of about 700 people on Facebook.

In her first interview, Priya spoke of how people used the WhatsApp group to get everyday information and support across issues as diverse as home baking, travel, and medical matters. When asked whether she had seen misleading information posted into the group, Priya reached for ways to explain. She explained that it was all about the rules: "there are rules in that group and there are about three admins, so if anything goes against their rules then it's, the person is deleted, or things like that. They have actions that are taken." Priya and her family had received help from the group on multiple occasions, and although they were technically "all strangers in a sense," she trusted them because they had to abide by the rules: "they have very strict rules, so it's all right, I guess," she said. Priya did not explain the rules' focus at that point, and the interview moved on to a different theme. In our analysis, we noted the use of passive voice in the phrase "they have actions that are taken."

Nine months later, Priya attended her second interview. After the first, she had voluntarily used our smartphone application to upload screen recordings of what she defined as useful or problematic information. Some were posts into the support group with "strict rules" she had mentioned in her first interview. She elaborated further on the group's rules, saying it was "basically against the rules of the group to discuss anything politics [*sic*] or religious or anything offensive."

The precariousness of these rules then became evident, and this case illustrates how the force of rules can subtly wither in larger groups. One of the materials Priya had uploaded to our project's database was a screenshot of a post by someone else in the group. This person had shared a link to a UK-based conspiracy theory "news" site. The

site features in some well-known fact-checker lists of misinformation sources, including Politifact and Health Feedback. The linked article contained several false tropes regarding how the UK National Health Service (NHS) presented Covid pandemic "excess deaths," for example, that changes in the way deaths were reported were a way to mislead the public. Importantly, Priya had categorized this example as "accurate and helpful" when she sent it to us. Had she accidentally uploaded it to the wrong category? we asked. No, "it seemed genuine then," she said. At the time, she had clicked on the link and read the article.

When the interviewer pointed out that the website was a reasonably well-known promoter of conspiracy theories, the fragility of the rules in this larger group was revealed.

Interviewer: "Do you think the person posting that link about Covid and the NHS, and the discussion, in your mind is that within the rules of the group as well?" [. . .]

Priya: "Yeah, uh. If it was a fake website, then it would have been flagged then, it would have been removed, but, uh, I didn't think it was a fake at that point in time, probably."

As the second interview drew to a close, Priya reflected on the moment she had clicked on the link to the conspiracy theory website's article: "I realized as we were speaking now [in the interview] and I went back to those posts, to think I should check the site before I clicked on it, so I need to be more careful."

Our aim here is not to admonish but to highlight the structural context in play. In this case, the relatively large community group comprising people with weaker ties had its own rules. In the past, these rules had played a role in enhancing trust among members. However, Priya also perceived such rules as something she did not play an active role in developing and applying—this was the responsibility of "three admins."[3] In the case of the link to the conspiracy theory website, no action was taken by the group's founders: it was not flagged and the poster remained in the group. In Priya's case, then, the context of the large group seemed to reduce the perceived need to engage in using the rules as a reference point to challenge misinformation. That was perceived to be the role of the people who first established the group and could control who could join. This seemed ultimately to diminish her agency. The rules had come to be perceived less as the consequence of vigilance among all group members and as more distant and delegated to others. She had clicked the link.

Conclusion

We used a qualitative research design and an interpretive analytical strategy to examine a previously hidden practice that can shape whether misinformation spreads online: group rules for personal messaging. We have discussed examples from family groups, a friendship group, neighborhood groups, a school group, a workplace group, and a large group in a city with a partially diasporic context.

Some of our participants developed rules to soften what they see as harmful affordances generated by the specificities of messaging as a hybrid public-interpersonal communication environment, particularly weak information provenance, minimal verification opportunities, and easy sharing. These affordances are perceived as increasing the circulation of misinformation that might potentially harm social ties by engendering misperceptions, sowing acrimony and division, or derailing the group's purpose. They are not determined by the features of the technology but emerge in the interactions between social and interpersonal relationships, patterns of use, and platform design (Evans et al., 2017).

In the cases we examined, rules were more powerful in groups comprising close ties. This is where personalized trust is more likely to operate, and it means misinformation inadvertently shared was also perceived to be more likely to deceive members of the group. Rules offer one way to reduce the resulting vulnerability.

We also found evidence that rulemaking can involve metacommunication that inculcates collective reflection and norms of epistemic vigilance (Sperber et al., 2010). Reflection on rules communicates social signals stressing the importance—as an end in itself—of developing desirable norms of discourse (Coe et al., 2014; Molina and Jennings, 2018). Rulemaking can be brief, informal, and minimal, but it can also be foundational. For instance, when a new group is first established, rulemaking can constitute a kind of "founding moment" with complex social implications, even if the task of creating a group on messaging platforms is trivial technologically. Rulemaking may also punctuate established groups at important moments, for example, when sharing information stimulates metacommunication about the norms of the group. Yet however fleeting or minimal it may be, rulemaking among these participants was a collective social process. It engaged people in reflection, with varying levels of intensity, on how social relationships and platform affordances can help misinformation spread and, importantly, how these forces can be blunted in routine, everyday interactions by rules. Founding moments of rulemaking can also have a long-term impact. This is because metacommunication signals a norm of ongoing vigilance and a normative assumption that rules are intrinsically valuable for mitigating the risk of harm.

Even if minimal and fleeting, this process had some positive long-term impacts among our participants. Our longitudinal research design revealed that some participants felt new rules reduced false and misleading information in the group over time. This can also boost perceptions that rulemaking mitigates misinformation's spread and it can serve to valorize rulemaking itself as a collective good. In our participants' larger groups with a mix of close and more distant ties, rules still led to norms of vigilance. However, the delegation effects of belonging to a larger group can diminish individual and collective agency and potentially leave members more vulnerable, as we saw in Priya's case. The perception can arise that the group's rules are the responsibility of others whose role is only vaguely understood.

In our fieldwork, rulemaking was not a particularly formal process but part of the spontaneous and self-organizing culture of personal messaging. Still, it involves people taking conscious decisions to keep misinformation and other harms out of the online spaces where they spend much time interacting with others central to their lives. Rules transcend the specific misinformation being shared and the resulting need to assign

blame to an individual for a specific act at a specific time. They momentarily shift matters to a more general plane of communication about the purpose of the group.

To be clear, our focus has been on rules that are informal *and* explicit; these are not unspoken or tacit "emotion rules" of the kind that Eliasoph (1998) identified as important in everyday conversations (pp. 235–236). At the same time, nor are they purely "procedural": they do not involve consideration of different decision rules such as the merits of unanimity versus simple majority voting or the assignment of speaking time, for example (Delli Carpini et al., 2004). The rulemaking we have identified is informal and not to be understood as the kind of "formal, rule-bound deliberation" that Eveland et al. (2011) contrast with what they call "informal political conversation." We see the practices we have unearthed as a distinctive subset of what Delli Carpini et al. (2004) termed "discursive participation" and not as public deliberation. Their role is primarily protective—to reduce misinformation harms to the group and its individual members—and is not geared to deliberately facilitating broader goals such as the formation of consensus or exposure to oppositional viewpoints. Our evidence suggests that, for some people, personal messaging enables this liminal orientation: these were cases where rules were made consciously and communicated explicitly, but the ethos of the encounters was oriented to care and protection and was not formally deliberative or procedure-focused.

Subject to further empirical research, these findings could have relevance for anti-misinformation initiatives across societies. Researchers and practitioners might further explore the impact of encouraging everyday rulemaking among ordinary citizens. While likely to have most impact in smaller groups, larger groups might also benefit, provided initiatives focus on developing agency and collective responsibility and account for the delegation of vigilance. An empathetic, dialogue-based approach is also more likely to be effective in larger groups, as the example of Sarah's neighborhood group illustrated.

Still, we must also contend with the reality that group rules clash with many attractions of online personal messaging, and this produces trade-offs. Rules are cumbersome. Some group members may perceive rulemaking as exclusionary forms of "censorship" or top-down attempts to exert power. Opportunities to challenge misinformation are lost once topics are placed off limits, as we saw with Lydia's experience, although it could be the case that rules operate in the background as a priming device, which warrants further investigation. It is relatively easy to exit a rule-bound group or form subgroups or one-to-ones, into which discussion of public affairs will inevitably flow, creating capillaries for misinformation. On the other hand, personal messaging is woven into everyday life and existing social and interpersonal relationships oriented around family, friends, workplaces, and localities. The embedded nature of groups in social life and the strong ties maintained within them likely constrain exits from the group (Baym, 2015; Matassi et al., 2019) especially when compared with public social media, where weak ties are more common.

That rulemaking on personal messaging platforms is decentralized, small scale, and informal is also an advantage. It is sometimes as much about the maintenance of harmonious personal relationships (Baym, 2015) as it is about developing critical awareness of misinformation. Among our participants, we found no evidence that rules were driven by grand notions of deliberative discourse. To encourage rulemaking, then, does not require the demanding optimism of many accounts of deliberation. Yet, if suitably encouraged,

rulemaking, even of an informal kind, could scale up and impact the flow of misinformation across larger networks. It can also go beyond the technological quick fixes much-vaunted by platform companies, such as restricting forwarding.

There is a broader context here. If we want to encourage rulemaking on personal messaging as one way to reduce the spread of misinformation, we might need to draw inspiration from earlier periods in the Internet's development, when community self-regulation through discursive practices was a more important organizing logic, before the dominant sociotechnical model (which is also a business model of course) of mass public social media platforms took hold. On public social media platforms, groups may coalesce around hashtags and follower/friend relationships. Selective following and use of blocking and "muting" features can, over time, help users curate something like group memberships. Platform algorithms, due to how they prioritize content, can also reinforce group belonging. Important as they are, however, these aspects of public social media do not depend much on conscious agency—on people coming together to collectively reflect on what kind of group they want to be and what kind of discourse they want to encourage. In short, because mass public social media platforms as currently configured rely so much on automated prioritization and sorting, they offer little inspiration or practical applicability when it comes to rulemaking on personal messaging. Instead, we need to learn from what ordinary people are doing on personal messaging and bring this into dialogue with themes from the research on online communities before social media. We need to bring people and their social capacities back in and build from there.

Our findings and the implications we discuss come with important caveats. First, our aim here is not to provide statistical frequencies generalizable to the population level; we have revealed themes that could be explored further using a range of different methods, including general population studies. Our method was not designed to recruit a fully representative sample of the UK population. Second, we reiterate that we examined a subsample of participants from a larger program of fieldwork. Rules or rulemaking was mentioned by about a third (33) of the overall participant cohort of 102 people. The topic of rules or rulemaking emerged organically when some participants referred to them while discussing their experiences of messaging. As a result, with our data we cannot draw conclusions about the statistical prevalence of these behaviors among the general public.

Further research on group rules and rulemaking on personal messaging is needed because these platforms have become extremely popular but create serious challenges for online anti-misinformation strategies. We argue these platforms require different, more contextual approaches based on how norms of interaction are established and maintained. Our analysis suggests that, while certainly not a panacea, encouraging group rulemaking might help reduce the spread of online misinformation.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is supported by a Leverhulme Trust Research Project Grant (RPG-2020-019).

ORCID iD

Andrew Chadwick https://orcid.org/0000-0002-5155-8173

Supplemental material

Supplemental material for this article is available online.

Notes

1. Lack of space precludes discussion of this research literature, but for a useful account see Dutton (1996).
2. The main exception is Telegram, which has a search function. We do not discuss it in this article due to its insignificant usage levels among our participants (see Supplementary Information Figure S1). However, we do include testimony from a Telegram user in the Supplementary Information file, Table S1.
3. A technical point here is that WhatsApp groups are set up by one user, who can then add or remove people. Of course, it was possible that the three admins to whom this participant referred worked together to make decisions, even if only one user had the technical ability to add and remove users.

References

Baym NK (2015) *Personal Connections in the Digital Age*. 2nd ed. Cambridge: Polity Press.

Bimber B and Gil de Zúñiga H (2020) The unedited public sphere. *New Media & Society* 22(4): 700–715.

Bird SE (2003) *The Audience in Everyday Life: Living in a Media World*. London: Routledge.

Boczkowski PJ, Suenzo F, Mitchelstein E, et al. (2021) From the barbecue to the sauna: a comparative account of the folding of media reception into the everyday life. *New Media & Society* 24: 2725–2742.

Coe K, Kenski K and Rains SA (2014) Online and uncivil? Patterns and determinants of incivility in newspaper website comments. *Journal of Communication* 64(4): 658–679.

Corbin JM and Strauss A (1990) Grounded theory research: procedures, canons, and evaluative criteria. *Qualitative Sociology* 13(1): 3–21.

Delli Carpini MX, Cook FL and Jacobs LR (2004) Public deliberation, discursive participation, and citizen engagement: a review of the empirical literature. *Annual Review of Political Science* 7: 315–344.

Dubois, E, Minaeian, S., Pquet-Labelle, A. and Beaudry, S (2020) Who to trust on social media: How opinion leaders and seekers avoid disinformation and echo chambers. *Social Media + Society*. https://doi.org/10.1177/2056305120913993

Dutton WH (1996) Network rules of order: regulating speech in public electronic fora. *Media, Culture & Society* 18(2): 269–290.

Eliasoph N (1998) *Avoiding Politics: How Americans Produce Apathy in Everyday Life*. Cambridge: Cambridge University Press.

Evans SK, Pearce KE, Vitak J, et al. (2017) Explicating affordances: a conceptual framework for understanding affordances in communication research. *Journal of Computer-Mediated Communication* 22(1): 35–52.

Eveland WP, Morey AC and Hutchens M (2011) Beyond deliberation: new directions for the study of informal political conversation from a communication perspective. *Journal of Communication* 61(6): 1082–1103.

Granovetter M (1978) Threshold models of collective behavior. *American Sociological Review* 83(6): 1420–1443.

Hameleers M and Van der Meer TGL (2020) Misinformation and polarization in a high-choice media environment: how effective are political fact-checkers? *Communication Research* 47(2): 227–250.

Hochschild AR (2016) *Strangers in Their Own Land: Anger and Mourning on the American Right*. New York: The New Press.

Kligler-Vilenchik N (2022) Collective social correction: addressing misinformation through group practices of information verification on WhatsApp. *Digital Journalism* 10(2): 300–318.

Ling R (2012) *Taken for Grantedness: The Embedding of Mobile Communication into Society*. Cambridge, MA: MIT Press.

McCroskey JC and Teven JJ (1999) Goodwill: a reexamination of the construct and its measurement. *Communication Monographs* 66(1): 90–103.

Malhotra P and Pearce KE (2022) Facing falsehoods: strategies for polite misinformation correction. *International Journal of Communication* 16: 2303–2324.

Masip P, Suau J, Ruiz-Caballero C, et al. (2021) News engagement on closed platforms. Human factors and technological affordances influencing exposure to news on WhatsApp. *Digital Journalism* 9(8): 1062–1084.

Matassi M, Boczkowski PJ and Mitchelstein E (2019) Domesticating WhatsApp: family, friends, work, and study in everyday communication. *New Media & Society* 21(10): 2183–2200.

Molina RC and Jennings FJ (2018) The role of civility and metacommunication in Facebook discussions. *Communication Studies* 69(1): 42–66.

Office of Communications (Ofcom) (2021) Online nation annual report. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0013/220414/online-nation-2021-report.pdf

Pang N and Woo YT (2020) What about WhatsApp? A systematic review of WhatsApp and its role in civic and political engagement. *First Monday* 25(1). Available at: https://firstmonday.org/ojs/index.php/fm/article/view/10417

Pasquetto IV, Jahani E, Atreja E, et al. (2022) Social debunking of misinformation on WhatsApp: the case for strong and in-group ties. *Proceedings of the ACM on Human-Computer Interaction* 6: 117.

Pearce KE and Malhotra P (2022) Inaccuracies and *Izzat*: channel affordances for the consideration of face in misinformation correction. *Journal of Computer-Mediated Communication* 27: zmac004.

Pink S and Mackley KL (2013) Saturated and situated: expanding the meaning of media in the routines of everyday life. *Media, Culture & Society* 35(6): 677–691.

Rossini P, Stromer-Galley J, Baptista EA, et al. (2021) Dysfunctional information sharing on WhatsApp and Facebook: the role of political talk, cross-cutting exposure and social corrections. *New Media & Society* 23(8): 2430–2451.

Saurwein F and Spencer-Smith C (2020) Combating disinformation on social media: multilevel governance and distributed accountability in Europe. *Digital Journalism* 8(6): 820–841.

Sperber D, Clément F, Heintz C, et al. (2010) Epistemic vigilance. *Mind & Language* 25(4): 359–393.

Swart J, Peters C and Broersma M (2019) Sharing and discussing news in private social media groups. *Digital Journalism* 7(2): 187–205.

Tandoc EC, Lim D and Ling R (2020) Diffusion of disinformation: how social media users respond to fake news and why. *Journalism* 21(3): 381–398.

Thorson K and Wells C (2016) Curated flows: a framework for mapping media exposure in the digital age. *Communication Theory* 26(3): 309–328.

Toff B and Nielsen RK (2022) How news feels: anticipated anxiety as a factor in news avoidance and a barrier to political engagement. *Political Communication* 39: 697–714.

Uslaner EM (2002) *The Moral Foundations of Trust*. Cambridge: Cambridge University Press.

Valeriani A and Vaccari C (2018) Political talk on mobile instant messaging services: a comparative analysis of Germany, Italy, and the UK. *Information, Communication & Society* 21(11): 1715–1731.

WhatsApp (2020) Two billion users. Available at: https://blog.whatsapp.com/two-billion-users-connecting-the-world-privately

Withagen R, De Poel HJ, Araújo D, et al. (2012) Affordances can invite behavior: reconsidering the relationship between affordances and agency. *New Ideas in Psychology* 30(2): 250–258.

Author biographies

Andrew Chadwick is Professor of Political Communication in the Department of Communication and Media at Loughborough University, where he also directs the Online Civic Culture Centre (O3C).

Natalie-Anne Hall is a Postdoctoral Research Associate for the Everyday Misinformation Project based in the Online Civic Culture Centre (O3C) in the Department of Communication and Media at Loughborough University.

Cristian Vaccari is Professor of Political Communication in the Department of Communication and Media at Loughborough University.