



# Sharing (mis) information on social networking sites. An exploration of the norms for distributing content authored by others

Lavinia Marin<sup>1,2</sup>

Accepted: 6 January 2021 / Published online: 2 February 2021  
© The Author(s) 2021

## Abstract

This article explores the norms that govern regular users' acts of sharing content on social networking sites. Many debates on how to counteract misinformation on Social Networking Sites focus on the epistemic norms of testimony, implicitly assuming that the users' acts of sharing should fall under the same norms as those for posting original content. I challenge this assumption by proposing a non-epistemic interpretation of (mis) information sharing on social networking sites which I construe as infrastructures for forms of life found online. Misinformation sharing belongs more in the realm of rumour spreading and gossiping rather than in the information-giving language games. However, the norms for sharing cannot be fixed in advance, as these emerge at the interaction between the platforms' explicit rules, local norms established by user practices, and a meta-norm of sociality. This unpredictability does not leave us with a normative void as an important user responsibility still remains, namely that of making the context of the sharing gesture explicit. If users will clarify how their gestures of sharing are meant to be interpreted by others, they will implicitly assume responsibility for possible misunderstandings based on omissions, and the harms of shared misinformation can be diminished.

**Keywords** Misinformation · Sharing · Posting · Social media · Social networking sites · Forms of life · Gossip · Rumour · Epistemic norms · Language games · Responsibility

## Introduction

Whether we knew it or not, we are all soldiers engaged in an 'arms race' against misinformation (O'Connor and Weatherall 2019, p. 176) and our daily acts of sharing or posting on social media influence this race. Whenever we open our social networking site of choice and we see an intriguing post, we are confronted with a flurry of micro-decisions: to share it further, to comment on it, to report it, or to ignore it altogether. Depending on our choice, the post can be propagated and thus made visible to our network of friends or stopped on its tracks. Our individual micro-decisions as users aggregate into a tsunami of information travelling on social media platforms, a flood increasingly polluted by misinformation and 'fake news'. Social networking sites such

as Twitter or Facebook are very efficient channels for the propagation of misinformation because of the massive informational content shared by their users – content that they themselves did not author but only shared further. Regular social media users are responsible for most of misinformation propagated on social networking sites (SNSs) (Jang 2018, p. 111; Nelson 2018, p. 3722) since misinformation would have much less harmful effect if it were not made visible by being shared by social media users. Without ignoring the effect that bots and for-profit propaganda sites have in creating and sharing misinformation, the role of regular users in amplifying the storm of misinformation deserves further scrutiny because their well-intended acts of sharing content have an aggregated disastrous effect on the online information ecosphere.

Sharing other's content is an everyday activity that most social media users partake in without much thought. Sharing happens as the result of a split-second decision, yet its effects are long-lasting and tend to ripple: since sharing amplifies misinformation to an unprecedented extent, it generates epistemic harms at collective and individual levels. The individual harm is that some people may acquire

✉ Lavinia Marin  
lavimarin@gmail.com

<sup>1</sup> Ethics and Philosophy of Technology Section, TU Delft, Delft, The Netherlands

<sup>2</sup> Education, Culture and Society, KU Leuven, Leuven, Belgium

misleading beliefs as result of seeing misinformation shared by their peers. The collective harm is that the general online ecosystem of information becomes polluted by misleading stories, and many users' energy and attention are diverted to non-issues created by fake news that may even end up in mainstream media as topics worthy of concern. Given these harmful consequences, an urgent question arises: Should there be some sort of accountability associated with the user's acts of sharing after all? Probably yes, but it is not that obvious what should these norms for sharing be. Many users still see Social Networking Sites as spaces for fun and leisure, and they do not expect to be held accountable to norms other than the basic norms regulating hate-speech and personal attacks which they agreed to when joining the platform. At first sight, it seems that there is less blameworthiness for merely sharing misinformation than for posting it, at least this is how many users justify their acts of sharing. When the content shared is proven to be misinformation, users will revert to justifications such as 'retweets are not endorsements' to showcase their well-meaning intentions. This normative ambiguity stems from the unclear nature of the norms for sharing: we do not know whether these norms are epistemic, moral, political, and aesthetic or none of these. Hence the aim of this article is to explore and clarify the kinds of norms that should govern regular users' actions of sharing content on social networking sites.

### On the ambiguity of norms for sharing on social networking sites

The rising tides of misinformation on SNSs constitute an important factor in destabilising democracy (Martens et al. 2018, p. 8) and an 'epistemic threat' (Goldman and O'Connor 2019) since democratic processes need informed citizenry for collective decision making (Mintz 2012; Goldman 2008, p. 111). The problem of misinformation on SNSs came to public attention after the US elections of 2016 (Hagood-Coote 2018, p. 23) when it was deemed that disinformation campaigns started online managed to change voters' minds just in time for the elections. Since then, various solutions have been implemented to deal with misinformation, usually relying on a combination of algorithmic approaches that combine machine learning techniques and human supervision (Zannettou et al. 2019, p. 14) such as user reporting, harvesting the problematic links and verifying new posts against a known database of hoaxes, or text mining of the posts for certain phrases. These semi-automated solutions propose a technological fix to what appears to be a normative problem: users do not seem to know or care when it is appropriate to share something for their networked connections. Misinformation is understood as any kind of 'misleading information' (Skrms 2010, p. 80) that is sent

around without an explicit aim to mislead others. A related term, disinformation, designates false information deliberately published with the intent to mislead (Skrms 2010, p. 80). Misinformation is a complex phenomenon that can take many formats—a text, a comment, an image, a video or sound-clip – and types: one can mislead by giving a false statement, but most often this happens by contextualising it wrong such inflating its significance or omitting important details (Wardle 2017). As Vosoughi and colleagues put it, on social media often 'falsity travels with greater velocity than the truth' (Vosoughi et al. 2018, p. 1149) and this has something to do with the ease with which one can share any kind of content created by others, with the click of a button. On SNSs, many regular users amplify misinformation by sharing it to their friends without a clear intention to mislead (Chen et al. 2015, p. 111), hence I prefer to use the term 'misinformation' to characterise any type of false content that is shared further by users. The destabilising effect of misinformation on SNSs is made visible because of the regular users' day-to-day acts of sharing that aggregate and propagate misinformation at exponential levels.<sup>1</sup>

There was no decisive moment when SNSs were invested with epistemic responsibilities by society. Many SNSs started as platforms for users to connect with their friends (Facebook<sup>2</sup>), to store and share publicly their own files (MySpace), or to voice their opinions uncensored to a self-selected audience of followers (Twitter). As SNSs became increasingly popular, an expectation emerged that SNSs should be held accountable for the misinformation shared on their platforms. This expectation seemed justified when news agencies created institutional pages on SNSs and used them to publicise news-content copied from their original websites. Some users who were not following these official pages were still exposed to the news after seeing the links posted by their friends. Nowadays, a large segment of the younger population has no source of news other than what their friends share on their preferred SNS (Wohn and Bowe 2016, p. 1). This means that, for some people, SNSs are their primary source of information about current events although they are informed by accident in a haphazard manner. Given the epistemic harms created by shared misinformation, it seems reasonable to expect that users should take some form of responsibility for any kind of content they share in their online social network. Yet it is unclear what norms are at stake for sharing since regular users are not journalists and

<sup>1</sup> Also see Temming, Maria. 2018. "On Twitter, the Lure of Fake News Is Stronger Than the Truth." *Science News*, March 3. Accessed November 12, 2019. <https://www.sciencenews.org/article/twitter-fake-news-truth>.

<sup>2</sup> Facebook's ancestor, 'Facemash', was intended as a website for rating the attractiveness of college students at Harvard, by comparing their pictures.

will refuse to abide by journalistic deontology since they perceive their acts of sharing as low-key gestures. This is seen in how users choose to eschew responsibility for consequences of what they shared such as the frequent disclaimer to be found on Twitter profiles stating that ‘retweets are not endorsements’. This creates the phenomenon of ‘bent testimony’ (Rini 2017) which is a kind of testimony without accountability, occurring when ‘[p]eople are happy to be understood as asserting the contents of shared news stories that turn out accurate (especially if they ‘scooped’ their friends) but insist that they meant no such assertion when trouble emerges’ (Rini 2017). For users, the norms for sharing are clearly different than those of posting, yet what these norms are is hard to tell.

Any attempt to clarify the norms for sharing needs to shed light on the very nature of the gesture of sharing and how this is different from posting of original content. Posting is the gesture of making public content created by oneself such as a text, a comment, a video, etc. Sharing is defined here as the behaviour of an online user transmitting further the information received from others with minimal alterations if any: this may include re-posting, re-tweeting, but also posting a link to a website where the misleading information was created (the post is not original because it merely signals the content placed elsewhere and written by others). At first sight, it seems that sharing should be governed by the same norms as posting, since both are accomplished by similar gestures and technological affordances: with the click of a button the content is made visible to one’s peers. Yet the gestures accomplished are different as these rely on different speech-acts. Posting is about asserting that something is the case and doing so in a public manner since posting is not speaking to oneself and someone else is bound to see the post. Meanwhile, sharing is a gesture of pointing – or, as Marsili (2020) has put it, “an act of ‘quoting by indicating’”. Thus, posting and sharing are different kinds of speech acts and the difference concerns the content of the assertion. If I post on my Facebook wall or Twitter page that “Donald Trump lost the 2020 elections”, I am asserting that it is true that D. T. lost the elections. If, however, I am sharing someone else’s post that “Donald Trump lost the 2020 elections”, I am pointing at the fact of the assertion being made and not at its content. The semantic content of my sharing is “Someone said that D. T lost the elections”. There are different conditions of truth for sharing which are easy to verify by backtracking the original post I shared. Yet since it is technically impossible to fake a share<sup>3</sup> because one cannot hide the source of the initial post, my assertion will always be true,

<sup>3</sup> This differs based on the platform used. A retweet or a share on Facebook can be edited, but the link to the original post remains so everyone can check by backtracking to the original source – hence it makes little sense to alter the words of others.

albeit a trivial truth. The pragmatic value of sharing is more than asserting that someone said something. To discern what is at stake in sharing, I side with Arielli (2018) who pointed out that sharing is a speech-act ‘whose aim is to direct the attention of other people to a content, stating (or expressing) its shareworthiness’ (Arielli 2018, p. 253). When we retweet or share a post, we are not stating that it is true, rather we are acting as conductors of our friends’ attention flow; our claim is that what we share is minimally interesting for others. A retweet would be nothing more than a gesture pointing at a piece of information with the subtext: “Look here, this is worthy of your attention”. In this interpretation of sharing as a gesture of pointing, the truth value of information shared is not the most important factor for the users. It is possible, of course, that I share something because it is true, but this is not clear if I give no explanation for the reasons of my sharing. Take the example of a Donald Trump tweet containing gibberish: “Despite the constant negative press covfefe” (Donald Trump 2017 cited in Marsili 2020). Marsili thinks that most of the 127 thousands retweets were ironic or pointing out its hilariousness, hence clearly not endorsements (Marsili 2020). Yet, without the context of each retweet, it is hard to conclude this because there are probably fans who retweet whatever Donald Trump tweets as signs of support, presumably not reading the original tweets. Without adding some sort of commentary on the retweet itself, we are not justified in assuming irony nor endorsement. Yet we can claim that the content is deemed minimally interesting since the users made the effort to share it. There is one important exception to this rule: if, when sharing a post, I also claim that I believe that it is true, then I am also asserting it. In these cases, where the context of sharing is clear and resembles testimony, I am responsible for asserting what I have shared under epistemic norms of truth-telling.

Concerning the norms for posting original content online, researchers will usually point at the epistemic norms of testimony. To remind the reader, testimonial knowledge is the kind of knowledge at which we cannot arrive ourselves, but that we indirectly know by trusting others (Lackey 2006, p. 3) be they experts in a field, eyewitnesses or people whose reasoning skills we trust. On SNSs, we find both experts who post with their real names, as well as eyewitness accounts from locations inaccessible to journalists. Without the testimonial trust in SNS users and their posts, the general public would have missed vital information about certain local events that began political revolutions such as the Arab Spring or Black Lives Matter. The relevance of testimonial norms for posting original content on SNSs is also seen by the fact that users should be held accountable for their statements and they can go to court for what they posted – as seen in several trials around Europe and the USA (Arielli 2018). To post something on a SNS is equal to asserting that it is true, and this implies legal liability at least. This does not

mean that the norms governing original posts are only epistemic, other norms can also apply, yet the epistemic norms are fundamental because to post something means to assert it publicly. Yet, if we look at sharing as a mere gesture of pointing at something as shareworthy, the problem of bent testimony morphs into a problem of unclear context: why are users sharing something that they do not endorse after all? They need to at least clarify their intentions and what they find shareworthy about that piece of shared content. If the context were clear, one could evaluate the act of sharing as testimonial or not. Yet fixing the context is hard to do beforehand and this seems to be a main challenge for most general SNSs. The reason why will be explained next.

### Forms of life and language games on social networking sites

SNSs do not offer one context of assertion because there is not one single or clear purpose served by SNSs. There is no such thing as one social network ‘to rule them all’, but rather a multiplicity of platforms allowing for user-generated content to be created and propagated in different forms. Often, the user norms of engagement are shaped by the particular purpose of the SNS. I will distinguish between types of SNSs to make it clear when misinformation arises as a problem of user-generated content and when it is a matter of user actions, particularly of sharing. This distinction categorises SNSs into either purposeful or general. *Purposeful SNSs* are social networks in which one main purpose aligns the user’s actions and gives the norms by which the users abide because the context of use is clear from the start. For example, informational SNSs such as question-and-answer (Q&A) sites are guided by norms that help users filter out useful from irrelevant information, and this includes sorting misinformation from genuine information. The mechanisms used are cooperation and communication, as users rate the good answers and flag the informational garbage. Other examples of purposeful SNSs are dating sites and job sites. The purpose on dating sites is to help users find their match. This process is tacitly understood by all users and enforced by them, for example the users will flag as inappropriate the trolls or the spammers. On jobs sites, the purpose is to establish showcase one’s employability and connect with employers; for this purpose, users will collaborate in endorsing each other’s skills and verifying their identity, or in sharing job ads that might interest their contacts. By contrast, *general SNSs* allow a variety of purposes, depending on what their users want to get out of their online social experience. Examples of general SNSs are Facebook and Twitter, not incidentally the platforms where misinformation is most rampant. On general SNSs, one can find any kinds of information, from news items posted by official agencies to cat pictures,

memes, and a flurry of users’ opinions about current events. General SNSs could be characterised as ‘online lifeworlds’ (Cocking and Hoven 2018, p. 35), different experiential realms with tacit rules of interaction. Here I will propose a related yet different conceptualisation for SNSs, namely of *online forms of life*—to use a Wittgensteinian term. This will allow me further to analyse the SNSs not in terms of the user’s experiences, but in terms of users’ practices—which are crucial in spreading misinformation.

When dealing with a purposeful SNS, the norms of use can be derived from its main purpose up to a large extent. Posting, endorsing a friend’s skills, rating someone else’s answer – are all actions that fall under norms of truth-telling. When I endorse a friend’s skills on LinkedIn, I am implicitly testifying that my friend has those skills. When I vote the answer to a user question on Stackexchange, I am testifying that I believe this answer to be appropriate for the question, and implicitly true. But since users can decide for themselves how to (mis) use a platform, this does not imply that we are guaranteed to find objective facts on purposeful SNSs. For example, a user group of anti-vaxxers posting on Quora can function as an ad-hoc epistemic community, with their own rules about what counts as evidence, even if their discussions will probably not produce knowledge relevant for the rest of the world. In the case of SNSs with an informational purpose, these need two additional conditions to function besides stating their informational purpose from the start: some technical scaffolding in the user interface and a community of users enforcing epistemic norms. On Wikipedia, editors do all the heavy work of checking other user’s contributions, while bots work non-stop to restore vandalised pages. On Q&A sites, the rating system is what makes possible to expect reliability: this is a technical scaffolding that, when used correctly, allows other users with expertise to endorse emerging experts or to encourage diligence and fact-checking by up-voting their answers.

By contrast, when users join a general SNS, their purpose does not have to be clear not even to them because SNSs allow for a potentially endless pool of purposes. Empirical studies from psychology and sociology have shown that there is not one single reason why users join and stay on a general SNS, but rather a dazzling multiplicity. For example, psychologists using the ‘uses and gratification’ model (Diddi and LaRose 2006; Lee and Ma 2012) claim to be ‘four main motivational categories: entertainment, socializing, information seeking, and self-expression and status seeking’ (Chen et al. 2015, p. 112) used to explain why users perform certain actions on SNSs and why they join these platforms in the first place. Other models are sociological, as found in Christian Fuchs’s work which identifies three main reasons for social interaction: ‘Information (cognition), communication and co-operation’ (Fuchs 2014, p. 42). Following Fuchs, if a SNS has an informational purpose, then the other two

modes of interaction—namely communication and cooperation—will be subservient to it. Other studies suggest that people use general SNSs for two major reasons: the ‘need to belong’ to a group and the ‘need for self-presentation’ (Nadkarni and Hofmann 2012). These empirical models and similar others indicate that there is no such thing as an over-arching narrative that would explain ultimately why people join and use SNSs, hence not a straightforward way to derive norms from this over-arching purpose. General SNSs remain purposeless exactly because they can accommodate an unpredictable variety of uses, yet these uses need to be agreed upon by groups of users. This agreement creates the local norms for users, emerging bottom-up from user practices.

Approaching SNSs as a collection of online forms of life allows us to look at the practices themselves without imposing any norms from the outside. In this view, describing the sharing practices means also describing the norms at the same time, as norms are immanent to practices: ‘the normative standards of use are immanent to use’ (Luntley 2003, p. 49). This has also the advantage that it does justice to the multiple ways in which SNSs have been used thus far, and to the diversity of SNSs out there. Following this approach, one should not speak of ‘misinformation on Social Media’, rather of misinformation in this or that context of practice, in this or that community of users. If SNSs are taken as showcasing forms of life, then sharing (mis) information is just another language game. As Wittgenstein had remarked, not all language games are about giving correct information: ‘Do not forget that a poem, although it is composed in the language of information, is not used in the language-game of giving information’ (Wittgenstein 1967, p. 160). Just like a poem is not about giving information, similarly the SNSs user’s activity of sharing news-worthy items by sharing or retweeting them on their pages is frequently not about informing others or testifying that these are true.

Misinformation can emerge from non-informational language games when it serves other purposes than knowledge-sharing. An empirical study looking into the motivations of users to share misinformation on SNSs pointed out overwhelmingly that ‘many of the top-ranked reasons were non-informational’ (Chen et al. 2015, p. 113), concluding that many ‘respondents share misinformation often for non-informational reasons’ (Chen et al. 2015, p. 114). These language-games are not specific only to SNSs. We sometimes say things we do not mean, out of politeness, and even if we mean the words, the utterance is not intended to inform others all the time. Lackey refers to these uses as ‘non-informational expressions of thought’ (Lackey 2006, p. 2) and opposes these explicitly to testimony. The non-informational expressions of thought may be just ‘conversational fillers’ (Lackey 2006, p. 2), things we say to promote group cohesion or to make the conversation continue further. Similar

to the offline life situations, not all content shared by users is meant to inform or to give grounds for the formation of beliefs. Information shared on general SNSs will not have a primary epistemic purpose unless the group of users agrees (even tacitly) that accurate information is their main goal. Thus, the structure of a Facebook closed group can be used to share educational materials among students and to inform each other about upcoming exams but it can also be used to share conspiracy theories behind the public scrutiny, if it is a group created by conspiracy fans. The community decides what is the purpose of their network, but also what counts as truth in their community and what counts as evidence.

If information giving is not the main nor the only reason why users distribute content on general SNSs, then it makes sense to look at other language games that could model closer to these practices. Starting from the observation that ‘catchiness rather than truthfulness often drives information (and misinformation) diffusion on social media’ (Chen et al. 2015, p. 111), I propose looking into rumours and gossiping as language-games that could function as models for the sharing of information online. The socialising functions of these ‘pathologies of testimony’ (Coady 2006, p. 253) seems to be the same, namely tied to ensuring group cohesion and trust among the selected group members. Gossip in particular has been proposed as an evolutionary mechanism that helped ensure group cohesion but also other functions such as “entertainment, cultural learning, sharing information, social bonding, and altering reputations” (Backer et al. 2016, p. 268). What appears as a ‘pathology of testimony’ to an epistemologist is in fact an everyday mechanism of socialising. This is not meant to say that rumours or gossip should be free of epistemic scrutiny, but rather that the context in which these language games appear is key to deciding if the stakes are about the information conveyed or the social relations.

When distributing gossip and rumours, the group cohesion is ensured by value allegiance as it helps us decide who shares the same values as us and who is an outsider – and this is easily seen by checking who relays our messages further on their SNS, who likes it and who comments it. Another function is virtue signalling (Rudnicki et al. 2019): two persons will gossip about a third to check if they both manifest the same attitude of disapproval or envy, etc. Often (mis) information shared on SNSs works similarly to gossip and rumour not because it is a secret information, but because it helps select and filter out who is on our group/network by approval and virtue signalling. This, in return, requires pragmatically that the user’s posts not be merely factual. Posting about the weather tomorrow will not work as a selection mechanism for whom to trust in our social network. We need to post reaction-seeking posts, emotional posts, moralising posts, i.e. post with which other users can agree or disagree in a way that shows their value allegiances.

Hence many users will be inclined to post and share content with a high emotional and normative charge, which implies that click-bait will be a likely candidate for sharing. From click-bait to misinformation sharing there is only one step left to take. This approach has the logical consequence that, just as sharing is not epistemic unless the context of the speech-act is clearly epistemic, similarly the posting of original content can be said to serve non-epistemic purposes. One can post disinformation just to test the value allegiances of one's followers and filter who is a faithful follower that shares the same worldview. However, with posting, since the post is an assertion, one can still be held accountable for fabricating disinformation even if one's purpose was to connect communities.

### Normative hierarchies on social networking sites and a meta-norm of sociality

Every SNS is made possible by a technical infrastructure that includes both physical tools as well as designed interfaces. My analysis thus far has focused on what individual users do and the aggregated effect of their individual choices: users can decide to use a SNS for its intended purpose or to misuse it, by altering its face altogether. But, as already pointed out by philosophers of technology, there is no such thing as a value-neutral technology (Winner 1980). Technologies always promote certain values to the detriment of others, values either embedded from the design phase, or emerging from their use (Friedman et al. 2013). If we were to inspect the infrastructures of general SNSs and thus try to infer what values these impose, we would find a variety of values, stemming from the general values promoted by ICT technology such as “human welfare, ownership and property, privacy, freedom from bias, universal usability, trust, autonomy, informed consent, accountability, identity, calmness, and environmental sustainability” (Huldtgren 2015). However, we should also pay attention to one specific meta-norm which emerges from users' practices on SNSs. This meta-norm is, I believe, what distinguishes SNSs from other digital platforms and ICT technologies.

The norms derived from the form of life are local and somewhat unpredictable, as these emerge from the ground-up, sometimes indifferent to the explicit purpose of the SNS. By contrast, there are also a few explicit norms in the Terms and Conditions agreements that forbid hate-speech and personal attacks but these are in place mostly for legal reasons and are kept to a minimum. Hate-speech and personal attacks are forbidden in the User Agreement and Terms and Conditions of most SNSs platforms; users agree with the Terms without usually reading them, but they still need to abide or else risk having their accounts suspended. Whenever a platform does not promote explicitly such Terms, it

is liable for lawsuits and its internet providers can be legally required to suspend their services towards such infringers. With the exception of these explicit norms, general SNSs do not impose explicit norms of behaviour to their users, rather guide their actions via one implicit meta-norm. This meta-norm makes possible that the user interactions take place, one could describe it as a condition of possibility for use. I will call it a *meta-norm of sociality* since it allows for the expansion of a user's network of connections. A concept introduced by sociologists in relation to norms of sanctioning (Horne 2001, p. 255), meta-norms explain why certain norms are enforced stricter on certain groups than on others, if enforced at all. Thus, first order norms are those we explicitly abide by in a group, and the meta-norms regulate to what extent the first order norms are enforced. Meta-norms are usually tacit and emerge through use. There is a significant possibility that other meta-norms are also at play in regulating the user interaction, but these are not discussed here.

Regardless of what the users' intentions are when joining a SNSs, they will be nudged to expand and connect with other users. This meta-norm is embedded in the technological affordances – what users can do and encouraged to – and in the designed interactions. For example, Facebook and Twitter explicitly encourage their users to increase the number of connections (Carr and Hayes 2015, p. 49) by suggesting new friends or possible acquaintances. SNSs by design allow their users to ‘view and traverse their list of connections and those made by others within the system’ (Boyd and Ellison 2007, p. 211) and this is difficult without a large number of connections. From this perspective, whatever content users post needs to abide by the local norms acceptable to the members of a group and, at the same time, to not hinder the meta-norm of sociality. An apparent violation of the meta-norm of sociality is the trolling behaviour—this consists in personal attacks, offensive language directed at a person or a group with the intent of making them leave the platform or the group altogether. The trolling works as an anti-social force that destroys social networks and diminishes the power of networks even though it enforces their own standing in the trolling community. By contrast, gossip, rumours, and urban myth-sharing can bond users and reinforces their ties, hence abide by this meta-norm of sociality. Sharing content posted by others helps to reinforces the networks of connections. On SNSs, users typically share other's content just as much as their own authored content (Lee and Ma 2012, p. 331). Since SNSs depend so much on the informational traffic generated by sharing, and since sharing is explicitly encouraged via the ‘share’ and ‘retweet’ buttons, this activity is seen as a valuable practice by the designers of SNSs. This may explain why one finds pieces of misinformation flying around the network long after these have been officially debunked (Friggeri et al. 2014) since the

discussions generated help the networks continue discussing and connecting.

To summarise, the following image shows the implicit hierarchy of norms on general SNSs that contributes to the general climate of normative ambiguity. On one hand, at the first layer, there are legal norms which are made explicit in the terms and conditions that all users agree to before joining the site. These norms are kept to a minimum and usually concern forbidding hate speech, personal attacks and other illegal activities. On the second layer, we have the meta-norm of sociality which makes sure that whatever users decide to do on SNS (within the boundaries of law) promotes the expansion of networks and socialising. Finally, the third layer, or the core, is made up of most norms that are local and unpredictable. Here we have the level where forms of life are manifested online, where language games are played, where groups derive (usually tacit) norms from practices. This third level is the one that generates most ambiguity since groups of users may decide to act as epistemic agents and share only trustworthy news, while other users may decide to have fun and share whatever they feel will get a reaction. If both tendencies generate traffic and expand the networks of users, while also not infringing any law through malicious content, both tendencies will be allowed by the SNS.

### On the users' responsibility for sharing

An important aspect of all SNSs, be those purposeful or general, is that these can be misused to a point of becoming unrecognisable. One could imagine a group of users starting to use LinkedIn as a dating site – contacting other users and asking questions about their profile that have nothing to do with their employability. Depending on how many users would engage in such behaviour, the SNS could be hijacked for other purposes or the users misbehaving could be banned. Thus, the norms of behaviour established by the platform owners may come at odds with the norms that groups of users choose to abide by. Given the locality of these norms, the norms for sharing also emerge locally. On a group of funny memes, my sharing of a picture means that I think it is funny. On a group of twentieth century art-lovers, my sharing of a painting means I think it is aesthetically valuable and shareworthy. If I mix up the groups and share a painting on the funny memes group, many members will think it is probably funny or that I intended it to be seen as funny. This is because of the implicit assumption that users will post only funny content. If I would want to share some serious content on this group, my sharing needs to be contextualised by some explanation placed in front of the text or image I share. Thus, an act of sharing without any additional

explanation usually follows the norms tacitly expected by the members of that online community.

Sharing in a group is different from sharing as an individual user. When I post a link to a site on my Facebook wall or tweet it, I am not sanctioned by any community of peers. The problem here arises from the heterogeneity of my audience: colleagues, friends, relatives and perfect strangers may see my shared link, depending who are my connections. Chances are high that my post will be misunderstood by some of these audiences. If I share a link to something of interest only to my colleagues, probably my family will not get it. If I share something funny that my niece has said, my colleagues will also not get it or may not care. A user sharing something on her wall will target indiscriminately a multiplicity of audiences or 'publics' (Pesch et al. 2020, p. 2216) who will probably be confused by the language game played here. It makes a difference if I share a funny meme in a closed group of friends where the repeated interaction created some tacit rules, than if I share it on my Facebook wall in a public manner. In the latter case, I should explain in a few words why I am posting it and what I believe about it, i.e. that it is funny. Without this minimal context, my sharing as gesture of pointing at someone else's content is bound to be confusing. A well-known approach to moral responsibility posits three conditions for it: a causal connection between the agent's actions and an outcome, the agent's knowledge of the consequences, and the agent's freedom to act (Noorman 2020). However, a pervasive problem with sharing on SNSs seems to be the unpredictability of consequences of our gesture. While we may intend to share something because it is "shareworthy", what is shareworthy may not be understood in the same way by our audience – depending on whom we perform this gesture for. And while we can never control how someone reads our retweet or shared post, we can at least make an effort to clarify the context of our sharing – if it is meant as a joke, as truth, as emotional expression, etc. By fixing this context we are, implicitly, making clear what language game we want to play and thus we submit ourselves to the judgement of our peers. To simply state that "retweets are not endorsements" without taking care of saying for each retweet what it means is bound to create miscommunication.

In the previous sections I have argued that there is not a single set of norms that governs users' gestures of sharing on SNSs. Rather, because of the multiplicity of practices emerging from the local uses, the users themselves decide what language games they are playing and thus judge other users against the emerging norms. From these observations, one could wonder whether the SNSs territory is a normless one where only the users decide what norms to follow and what counts as truth. I do not wish to claim this, since the online realm does intersect often with the offline life and misinformation shared online has actual consequences in

people's lives. After all, it is almost impossible to separate the offline from the online worlds, since even people who refuse to use the Internet are affected in their daily lives by how others decide to behave online. There should be norms for sharing as there are for posting but, in order to make these explicit, we need to look at three different factors: the practices embedded in the form of life, the technological infrastructure, and the context of the individual gesture of sharing.

The difficulty with pinpointing the norms surrounding sharing on SNSs stems from a conjunction of factors at play. First, the technical infrastructure of general SNSs is privately owned by companies. This infrastructure changes all the time as platform designers keep experimenting with new designs and modes of engagement – see for example how Facebook makes visible design changes every year, while minor design tweaks are released silently all the time.<sup>4</sup> On SNSs, most users are guests with little authority over what norms are enforced, while the platform owners can decide how and if to enforce any explicit norms. This makes it that communities of users need to comply with the explicit norms of the platform providers, or leave the SNS altogether. However, if the platform providers are the only de facto enforcers of norms, this puts them in a conflict of interests since they aim for enforcing the meta-norm of sociality above all else. Whatever makes the users engage with others and expand their networks is good and will be usually allowed to take place, even if it infringes the first-order norms. In other words, gossiping and rumour mongering will be tolerated and encouraged since these expand and re-enforce connections. Misinformation sharing works in very similar ways with rumours and gossip. Second, there are local norms emerging from the language games played by different communities of users in groups or purposeful SNSs. These norms cannot be in conflict with the explicit norms of the platform or with the meta-norms (levels 1 and 2 in Fig. 1), but otherwise almost anything is permitted – provided that the first two levels of norms are not infringed. Finally, the user's gestures of sharing something publicly creates conflicting understandings since we do not know who will actually see our sharing gestures and if this is the intended audience after all. To summarise, the difficulty in fixing the norms for sharing on SNS lies in the conflict between values embedded in the technical infrastructure, the unpredictability of local norms, and the unfixed context of the gesture of sharing when the audience is no one in particular.

There is no easy solution to the problem of local and user-generated norms for sharing. One could even say

<sup>4</sup> This article illustrates how the Facebook profile and newsfeed have radically changed in a matter of 10 year: <https://www.dailydot.com/debug/old-facebook-profiles-news-feeds/>.

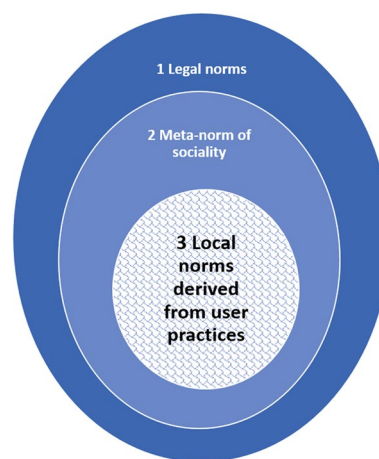


Fig. 1 Hierarchy of norms on SNS

that the same problems emerge also for posting original content, not merely for sharing. Yet, when I post a text written by me on a general SNS, I am asserting it and its context may be inferred from the content itself. Meanwhile, the gesture of sharing is much more obscure. It is a gesture of pointing at something with unclear intentions. Thus, when positing norms for sharing, at least the demand to make the context clear should be non-controversial. The clarification of context thus emerges as a basic user responsibility when sharing content in a public manner. The technical infrastructure of sharing can be changed provide contextualisation by default by asking users to say something about what they share before they do it. Twitter experimented with this feature in the wake of the 2020 USA elections.<sup>5</sup> However, the lack of a technical nudge should not let users off the hook. If they do not explain what they think about what they share, their post will be misinterpreted. Thus, a simple explanation even in one word – funny, interesting, outrageous, “true dat” – can help the other users grasp whether I am sharing information or just pointing at things that gave rise to an emotional reaction in me. In cases where context is made explicit, misinformation shared is not necessarily an irresponsible act. If I share misinformation to show

<sup>5</sup> Starting today, when you click or tap the retweet icon, Twitter will pull up the Quote Tweet composer to encourage you to write something about that tweet before you share it. You don't have to write anything if you don't want—just leave the composer blank and hit the retweet button to retweet like you normally would. But Twitter is hoping that by introducing some friction into the process, people might better consider exactly what they're retweeting or take the opportunity to add their own perspective. (source: <https://www.theverge.com/21524092/twitter-temporarily-changing-retweet-quote-tweet-election>).



my friends how outrageous I find these claims, or if I share it because I find it funny, I am also expressing quite clearly that I do not think it is true. But even if I share misinformation that I deem to be true, I will be less blameworthy than if I do it without any explanation and then, when confronted, back off while saying I did not mean it. Sharing misinformation is not in itself a blameworthy practice if users make the effort of explaining each time what was their intent. Even someone sharing a conspiracy theory and stating that they believe it is not causing necessarily an epistemic harm. Their endorsement of a conspiracy theory is a personal belief that they share with the world and, thus, they are playing an information-giving language game. In this case, they must be responsive to requests from other users for further reason-giving, for examining the authority of the link or for additional evidence. If they think what they share is the truth, then they must also be willing and able to enter into debates about this truth. Misinformation shared in this manner is what starts a dialogue and thus has the option of being publicly debunked by others. We can contrast this with the more harmful behaviour of sharing a link to a conspiracy site, no explanation given and then refusing to engage any demand for explanation from other users since it was not intended for them. In this latter case, the user is acting as if she is talking to herself, posting things in a public space as if it were private, gesturing for oneself. Any act of my sharing of a belief based on a misinforming link is informative in another way for my friends: it shows them that this is where I stand and, if they have any expertise in that topic, they can choose to confront me, educate me, bring counter-arguments; or my friends may choose to silently ignore me when I share news about that topic, as they classify me as untrustworthy in that regard. If, however, my friends are just as ignorant as me about this topic, they should not get their information from my shares and posts since I have not been established as an expert in that topic; hence their epistemic vice of laziness and not seeking additional information should not be blamed on my sharing my ignorance on a social network. It is almost never a virtuous epistemic behaviour to take one's information from one source (Worsnip 2019), especially if that source is some random user on a SNS.

## Conclusions

General SNSs such as Facebook and Twitter are infrastructures designed to facilitate the social networking of their users and expanding their networks. At some point, the demand emerged

that this medium of pure sociality also function as a medium of truth-telling, as a reaction to the epistemic harms created by misinformation posting and sharing. Yet if we were to ask general SNSs to promote solely truth-telling among their users, we would be confusing language games of truth-telling or information giving with other language games aimed at social cohesion, such as rumour sharing or gossiping. Since not all information is meant to inform others (Wittgenstein 1967, p. 160), conversely, not all misinformation shared on SNSs is meant to mislead. Misinformation sharing seems to belong more in the realm of rumour spreading and storytelling rather than in the information-giving game. Whenever we share something, we are acting in a kind of public space where anyone can see and misinterpret our gesture, hence as a minimal act of responsibility, we need to pre-emptively clarify what we mean with our gestures. Even if we do not intend to inform others, it is our epistemic responsibility to make clear that we are not playing an information-giving game. Information-giving language games are possible precisely because non-informational language games coexist; this diversity of language uses and gestures is a strength of social networking platforms and something to be celebrated. However, this observation does not leave us with a normative void: there are user responsibilities to be upheld whenever sharing content, be that information as well as misinformation. If one shares a link in a thematic group or in a purposeful SNS, the utterance will be judged against the norms of use emerging in that context, i.e. against the background of a certain form of life. However, if one relays (mis)information out in the open, by posting it on one's Facebook wall or tweeting it to all followers, then the context needs to be made explicit as well as the intention: is this meant to be taken as true information, as something funny, as an expression of outrage? Sharing is a pointing gesture whereby the content pointed at is deemed "shareworthy" (Arielli 2018) yet what is worthy of being shared depends on one's personal preferences as well as the intended audience. Without making clear the context with every gesture of sharing, the meaning of the gesture will become a source of confusion and can only aggravate the already destructive effects of online misinformation.

**Acknowledgements** I would like to thank Janna van Grunsven for giving inspiring feedback on the first version of the manuscript and to Mark Thomas Young for carefully reading the revised version. The following members of the audience contributed valuable insights during the presentations of this manuscript at various conferences and colloquia: Udo Pesch, Sabine Roeser, Andrea Gammon, Samantha Copeland, Martin Sand, Maarten Franssen, Pieter Vermaas, Lieke Brackel, Matthew Dennis, Owen C. King, Konrad Rudnicki.

**Funding** This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant agreement No 707404. The opinions expressed in this document reflect only the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Alvin, G. (2008). The social epistemology of blogging. In J. van den Hoven & J. Weckert (Eds.), *Information technology and moral philosophy* (pp. 111–122). Cambridge: Cambridge University Press.
- Alvin, G., & O'Connor, C. (2019). Social epistemology. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy, fall 2019: Metaphysics Research Lab*. Stanford, California: Stanford University.
- Arielli, E. (2018). Sharing as speech act. *Versus*, 47(2), 243–258.
- Backer, C. J., Larson, C., Fisher, M. L., McAndrew, F. T., & Rudnicki, K. (2016). When strangers start to gossip: Investigating the effect of gossip on cooperation in a prisoner's dilemma game. *Evolutionary Psychological Science*, 2(4), 268–277. <https://doi.org/10.1007/s40806-016-0063-7>.
- Boy, D., Danah, M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210–230. <https://doi.org/10.1111/j.1083-6101.2007.00393.x>.
- Carr, C. T., & Hayes, R. A. (2015). Social media: Defining, developing, and divining. *Atlantic Journal of Communication*, 23(1), 46–65. <https://doi.org/10.1080/15456870.2015.972282>.
- Chen, X., Sin, S. C. J., Theng, Y. L., Lee, C. S. (2015). "Why do social media users share misinformation?" In: PL Bogen, S Allard, H Mercer, M Beck, SJ Cunningham, D Goh, G Henry (Eds) *Proceedings of the 15th ACM/IEEE-CE on Joint Conference on Digital Libraries - JCDL '15*, ACM Press, New York, New York, USA, pp. 111–14
- Coady, C. A. J. (2006). Pathologies of testimony. In J. Lackey & E. Sosa (Eds.), *The epistemology of testimony* (pp. 253–271). Oxford: Clarendon.
- Cocking, D., & van den Hoven, J. (2018). *Evil online*. Blackwell public philosophy. Hoboken: Wiley.
- Diddi, A., & LaRose, R. (2006). Getting hooked on news: Uses and gratifications and the formation of news habits among college students in an internet environment. *Journal of Broadcasting and Electronic Media*, 50(2), 193–210. [https://doi.org/10.1207/s15506878jobem5002\\_2](https://doi.org/10.1207/s15506878jobem5002_2).
- Friedman, B., Kahn, P. H., Borning, A., & Huldgtren, A. (2013). Value sensitive design and information systems. *Early engagement and new technologies: Opening up the laboratory* (pp. 55–95). Dordrecht: Springer.
- Friggeri, A., Adamic, L. J., Eckles, D., Cheng, J. (2014). "Rumor cascades." In *Eighth International AAAI Conference on Weblogs and Social Media*.
- Fuchs, C. (2014). *Social media: A critical introduction*. London: SAGE.
- Habgood-Coote, J. (2018). Stop talking about fake news! *Inquiry*, 8, 1–33. <https://doi.org/10.1080/0020174X.2018.1508363>.
- Huldgtren, A. (2015). Design for values in ICT information and communication technologies. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Eds.), *Handbook of ethics, values, and technological design: Sources, theory, values and application domains* (pp. 739–767). Dordrecht: Springer Netherlands.
- Jang, S. M., Geng, T., Li, J.-Y., Xia, R., Huang, C.-T., Kim, H., & Tang, J. (2018). A computational approach for examining the roots and spreading patterns of fake news: Evolution tree analysis. *Computers in Human Behavior*, 84, 103–113. <https://doi.org/10.1016/j.chb.2018.02.032>.
- Lackey, J. (2006). Introduction. In J. Lackey & E. Sosa (Eds.), *The epistemology of testimony* (pp. 1–21). Oxford: Clarendon.
- Lee, C. S., & Ma, L. (2012). News sharing in social media: The effect of gratifications and prior experience. *Computers in Human Behavior*, 28(2), 331–339. <https://doi.org/10.1016/j.chb.2011.10.002>.
- Luntley, M. (2003). *Wittgenstein: Meaning and judgement*. Oxford: Blackwell.
- Marsili, N. (2020). Retweeting: Its linguistic and epistemic value. *Synthese*. <https://doi.org/10.1007/s11229-020-02731-y>.
- Martens, B., Aguiar, L., Gomez-Herrera, E., Mueller-Langer, F. (2018). "The digital transformation of news media and the rise of disinformation and fake news: JRC digital economy working paper 2018-02." Retrieved December 1, 2019 from <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/digital-transformation-news-media-and-rise-disinformation-and-fake-news>.
- Mintz, A. P. (Ed.). (2012). *Web of deceit: Misinformation and manipulation in the age of social media*. Medford, NJ: CyberAge Books.
- Nadkarni, A., & Hofmann, S. G. (2012). Why do people use Facebook? *Personality and individual differences*, 52(3), 243–249. <https://doi.org/10.1016/j.paid.2011.11.007>.
- O'Connor, C., & Weatherall, J. O. (2019). *The misinformation age: How false beliefs spread*. New Haven: Yale University Press.
- Pesch, U., Huijts, N. M. A., Bombaerts, G., Doorn, N., & Hunka, A. (2020). Creating "Local Publics": Responsibility and involvement in decision-making on technologies with local impacts. *Science and engineering ethics*, 26, 2215–2234. <https://doi.org/10.1007/s11948-020-00199-0>.
- Regina, R. (2017). Fake news and partisan epistemology. *Kennedy Institute of Ethics Journal*, 27(2S), E43–E64. <https://doi.org/10.1353/ken.2017.0025>.
- Rudnicki, K., de Backer, C. J. S., & Declerck, C. (2019). The effects of celebrity gossip on trust are moderated by prosociality of the gossipers. *Personality and individual differences*, 143, 42–46. <https://doi.org/10.1016/j.paid.2019.02.010>.
- Skyrms, B. (2010). *Signals: Evolution, learning, and information*. Oxford: Oxford University Press.
- Soroush, V., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>.
- Wardle, C. (2017). "Fake news. It's complicated." Retrieved 23Aug 2019 from <https://firstdraftnews.org/fake-news-complicated/>.
- Winner, L. (1980). Do artifacts have politics? *Daedalus*, 1, 121–136.
- Wittgenstein, L. (1967). "Zettel: GEM Anscombe, GH Von Wright (Eds)", Berkeley, CA: University of California.
- Wohn, D. Y., & Bowe, B. J. (2016). Micro agenda setters: The effect of social media on young adults' exposure to and attitude toward news. *Social + Media Society*, 2(1), 205630511562675. <https://doi.org/10.1177/2056305115626750>.
- Worsnip, A. (2019). The obligation to diversify one's sources. In C. Fox & J. Saunders (Eds.), *Media ethics, free speech, and the requirements of democracy, Routledge research in applied ethics 13* (pp. 240–264). New York: Routledge.
- Zannettou, S., Sirivianos, M., Blackburn, J., & Kourtellis, N. (2019). The web of false information. *Journal of Data and Information Quality*, 11(3), 1–37. <https://doi.org/10.1145/3309699>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.