

## Sharing Content Online

Alex Davies

### **1 Context-Sensitivity and Content-Sharing**

Context shapes linguistic content.<sup>1</sup> When George Grylls (2022) told his British readers that Estonia had requested more NATO troops to be sent to “the Russian border,” he was using the construction to refer to three sovereign countries: Estonia, Latvia and Lithuania. But when a visitor to the Estonian town Narva tells her host that she’s going to “the Russian border”, she’s not referring to three sovereign countries. She’s already there. She’s referring rather to something like a small checkpoint between the Republic of Estonia and the Russian Federation. The size of the referent of “the Russian border” waxes and wanes with the construction’s context of use.

Context shapes the content of more than just definite descriptions: it shapes the strength of modal verbs (which range of possibilities are you considering when you suggest “it might rain”?) and habituals (we’ll all lie in *some* circumstances, so which range of circumstances are you saying Edith will lie in when you say “Edith lies”?), the cut-offs and scales of gradable adjectives (in what way, and to what extent, must a person be proficient for the accolade “intelligent” to be applicable to them?), the things counted by count nouns (what counts as a city, a philosopher, an Irishman?), the stuff of mass nouns (do I give you water when I give you tea?), the size of the referent of an indexical (how big is “here”? how long is “now”?), and much else besides.

If this much linguistic content is context-sensitive, then we generally cannot assume that shared knowledge of a common language suffices for two people to ascribe the same content to the same construction (i.e. share content). Whether it’s likely or unlikely that two people will succeed in this is going to depend greatly on the environment of those two people. Moreover, it seems *prima facie* unlikely that all environments will be alike in how well they support content-sharing within them. It seems much more likely that some environments will be more conducive to content-sharing than others: some will make it likely that two people will ascribe the same content to the same construction and some will make it unlikely.<sup>2</sup> To take a simple example, suppose the local clientele of a restaurant in a countryside village doesn’t include any vegetarians and moreover that, in this village, there is a

---

<sup>1</sup> In this paper, I distinguish between content (the contribution a construction makes to the proposition a declarative sentence containing the construction literally expresses) and standing-meaning (the context-invariant meaning of the construction). With the exception of my remarks in section 2 on the work of (Frost-Arnold, 2021) and (Record & Miller, 2022) (where I mean to refer also to the pragmatic phenomena they discuss), when I refer to context-sensitive content, I mean one of two things: either the standing meaning of the construction underdetermines the construction’s content, or, the construction is polysemous i.e. has multiple related standing meanings. In either case, the underdetermination of the construction’s content can be resolved by context. In either case, the contextually resolved content that I mean to speak of will be part of the content that is taken to be said or asserted in uttering the declarative sentence.

<sup>2</sup> There are those who think that content-sharing with context-sensitive language basically never happens (e.g. (Abreu Zavaleta, 2019), (Bowker, 2019), (Peet, 2016)). For criticism of some of the arguments for this view see: (Davies, 2019) and (Davies, 2021)). These theorists can nonetheless allow that environments vary in their support of content-sharing by distinguishing between different degrees of success in content-sharing (an observation I owe to Joey Pollock (pc)).

commonly held view that a meal without meat is incomplete. The restaurant is under no economic pressure to design their signs and labels to facilitate vegetarians' decision-making when ordering. For these reasons, the restaurant puts chunks of pork in what it calls the "vegetable soup." Nonetheless, from time to time, vegetarians do visit from out of town. As a result, when they do, there is a regular failure to share content with the expression "vegetable soup." This contrasts with a restaurant environment in which the size of the vegetarian clientele provides a far stronger economic incentive to override any local mores. In such an environment, the regular failure to share content that we witness in the village will be absent.

## **2 Social Media Platforms and Bystander Information**

Notoriously, social media platforms are environments where content-sharing is attempted. So, given what was said in section 1, we can ask: are there distinctive features of social media platforms which make them especially inhospitable for the sharing of content using context-sensitive language?

One such distinctive feature—one which has already received scholarly attention—is the ease with which a vehicle of content that has been designed to convey content to one audience can end up getting consumed by a distinct and unanticipated audience ((Frost-Arnold, 2021), (Record & Miller, 2022)). You post a text to one network of people, for whom you have designed the text so that it's likely to be understood (by them) as you intend. But with the click of a "share" (or, depending on the algorithms in play, even a "like") button, someone else can take your text and present it to a different network of people, for whom you have *not* designed the text. If you used context-sensitive language in the post, then there's a significant risk that your post will be understood by the unanticipated audience to bear a content you didn't intend: a vehicle of content was shared, but the intended content was not. The same kind of risk is not nearly so great for other communication environments (quiet chats, private emails with people you know well enough, etc.) because the risk of exposure to unanticipated audiences is much lower.

This paper takes a look at a second distinctive feature of social media platforms, which has the potential to disrupt the sharing of content thereon. Like the first, it does so not by disrupting the sharing of vehicles of content (as with "filter bubbles" (Pariser, 2011)), but by disrupting the content that those vehicles are perceived to bear.

In contrast with dominant 20<sup>th</sup> century news media, posts on social media are typically and saliently accompanied by what I'll call "bystander information": information about the attitudes of other readers of a post (in the form of reactions and comments).<sup>3</sup> The quintessential solitary reader reads privately. They engage in a quiet exchange with the author. They mull over what they read without taking a public stand on either their understanding of what was said or on their attitude toward what was said. In contrast with the quintessential, social media socializes reading: there, the reader reads in the inescapable knowledge of what others' attitudes towards the same post are; and in "reacting" to a post, the reader's judgments-in-reading exhibit the character of a public reflex (like a grimace or an

---

<sup>3</sup> Letters to the editor have never been as salient as bystander information on social media and they have always been curated by those affiliated with the authors of the text (Chung, 2017, p. 950). They can be curated in order not to disturb the intended content of that text. Not so on social media.

unguarded cock of the eye in full view of the reflex's object). This all takes place, so to speak, *inside* the activity of reading: before the reader has finished making some kind of sense of what they are reading and before they have settled on a judgment about it (whether they agree with it). That's different from what might happen after the reader has digested what they have read when, of course, they *will* be likely to discuss what they have read with others. Those who don't like their second-hand books plastered with the judgments of previous readers don't, for all that, refuse to talk about the book with others after they've read it. It's the formative process of reading they want to keep clean of outside influence and it's precisely the privacy of this formative process that social media up-ends.

One effect that one might think bystander information has on consumers of social media is that it changes minds about a fixed object of judgement: the reader agrees with a fixed content of a post if told others agree with the post and disagrees with that very same content if told that others disagree with the post. Indeed, researchers working on the effects of bystander information by and large take this view about bystander information's influence (as we'll see in section 3). However, this is not the only explanation of the evidence. A second possibility is that bystander information changes what social media consumers believe a post's content to be.<sup>4</sup> The reader takes the post to have one content if told that others agree with it (a content they already agree with) but takes the post to have a different content if told that others disagree with it (a content they already disagree with). At no point does the reader change their attitude toward any one content. The reader rather changes which content they think they're being asked their attitude about. If bystander information can have this kind of effect, then social media might well complicate the sharing of content on social media in a way that has yet to be appreciated.<sup>5</sup>

This paper gives detail to this second possibility and makes a case for taking it seriously—worthy of further investigation. I begin, in section 3, by sketching the evidence collected to date on the effects of bystander information on consumers of social media. As I said in the previous paragraph: those who collect this evidence assume that it supports the conclusion that bystander information changes attitudes towards constant objects of judgment. However, in section 4, I present evidence of the capacity of bystander information to change the perceived *content* of linguistic stimuli. I explain how this capacity undermines the inference from the evidence presented in section 3 to the conclusion that bystander information changes attitudes towards a constant object of judgment. In section 5, I present three mechanisms (suggested by the evidence considered in section 4) by means of which bystander information could generate changes in the object of judgment. I use these mechanisms to identify two sets of risk factors that make such changes more likely. In section 6, I explore how, if active in the context of social media, these mechanisms have the potential to surreptitiously undermine the sharing of content (despite successful sharing of the vehicles of content: viz. posts). Finally, in section 7, I make clear that, if social media does make trouble for content-sharing in the way here envisioned, it will not be because of the inclusion of bystander information alone: it will be because social media injects such

---

<sup>4</sup> Although I'm speaking about what social media consumers believe the content of a post to be, I do not mean to pick a side in the debate on whether understanding what is said requires having a belief about what is said or instead merely entertaining what is said (Longworth, 2018).

<sup>5</sup> The present paper focuses on how social media structures linguistic interpretation. In this it contrasts with Nguyen's (2021), which focuses on how social media structures linguistic production.

information into the act of reading while at the same time failing to introduce factors that wash out that information's splintering influence.

### 3 Bystander Information and Attitudes: Some Evidence

One method for studying the effects of bystander information on attitudes towards a post is to conduct an experiment with the following structure. There are two conditions: in one condition, a post is presented with positive bystander information (i.e. bystanders agreed with the post); in another condition, a post is presented with negative bystander information (i.e. bystanders disagree with the post). Participants are randomly assigned to one of the two conditions. Their attitudes toward the post are elicited. One then checks for differences in attitude between the two conditions.

Studies with this structure have found differences in attitudes across the two conditions: for both comments ((Dixon, 2020), (Lee, Jang, & Chung, 2021), (Zerback & Töpfl, 2022), (Wijenayake, Hettiachchi, Hosio, Kostakos, & Goncalves, 2021)) and reactions (Lee, Atkinson, & Sung, 2022). More specifically, there's apparent conformity: people's attitudes are coincident with the predominant attitude expressed in reactions or comments. That's why, in this literature, the effect of bystander information on attitudes is sometimes called the "bandwagon effect" (following (Simon, 1954) and (Sundar, 2008)).

Such effects are often mediated by other variables. In particular: with reactions, this apparent conformity has been found to arise only if:

- you take reactions to be indicative of broader opinion (Chung, 2019)<sup>6</sup>
- you have a fear of social isolation (Porten-Cheé & Christiane, 2020)
- you like engaging in effortful cognition (Lee & Jang, 2010)

With comments, this apparent conformity has been found to arise only if:

- comment argument quality is high (Winter, Krämer, & Liang, 2017)
- you have a political leaning that correlates with preference for a certain kind of evidence (viz. anecdotal or scientific) and the comment presents your preferred kind of evidence (Hinnant, Subramanian, & Young, 2016)
- the source is a high credibility news organization (Chung, 2017)
- you like engaging in effortful cognition and you are confident that you are competent to understand, and participate effectively, in politics (Lee J. , 2014)
- you have commented yourself (Sah & Peng, 2022)
- the comments are casting doubt on the original post (Kluck, Schaewitz, & Krämer, 2019)

I'm being circumspect in saying "*apparent* conformity". The conclusion drawn by researchers in this literature is that we are witnessing causation on attitudes: i.e. that we're seeing bystander information causing shifts in attitudes. For reasons hinted at in section 2, and expanded upon in the next section, I think the present circumspection is required.

---

<sup>6</sup> Some find that this is not necessary for an effect stemming from comments (Winter, Brückner, & Krämer, 2015). Others find that it is ((Kim Y. , 2015) and (Kim, Han, & Seo, 2020)).

#### **4 The Asch Critique: Bystander Influence on Linguistic Interpretation**

Classic experiments in social psychology that elicit apparent attitudinal conformity are the stuff of lore (Asch, 1951). Three lines (A, B, and C) are presented to a participant sitting in a room with others attending to the same lines. A fourth line (D) is presented and all present are asked which of A-C is most similar in length to D. Unbeknownst to the participant, the others in the room are stooges who have been asked to act as though a line that is obviously different in length from D is the right answer. After the stooges have been asked their judgment, the participant is then asked for their own judgment. The participant very often gives the same answer as the stooges. A similar experiment generates similar results: participants are placed in booths ((Crutchfield, 1955) – a “Crutchfield situation”). In the booth is a light-board, which supposedly indicates the judgments of other participants in other booths in the same room. Each participant’s light-board gives each participant the misleading impression that all other participants have picked a line that obviously doesn’t match the length of the fourth line. Apparent conformity ensues. These apparent conformity effects have been replicated since ((Bond & Smith, 1996), (Bond, 2005)).

One explanation of what’s going on is that participants are losing confidence in their belief about the relevant line lengths because of the bystander information presented: i.e. they have an understanding of the question they’ve been posed that remains constant through the presentation of bystander information. The only thing that changes is their attitude toward this constant object of judgment. They follow the bandwagon.

Another explanation treats as pliable that which the first explanation treats as invariant: the behaviour of the participants is explained by a change in their understanding of the question they’ve been posed—the object of judgment—and their attitudes toward different objects of judgment do not shift. For instance, they start to doubt that they’ve understood what they’re supposed to do in response to the lines rather than losing confidence in which line is equivalent in length to D.

The latter change-in-the-object-of-judgment explanation was one favoured and promoted by Solomon Asch himself ((Asch, 1940), (Asch, 1948)). He found evidence of a change in the object-of-judgment when using a linguistic stimulus (as opposed to the lines of Asch (1951)). Tell people that most people dislike politicians, and they’ll tell you they dislike politicians too. Tell people that most people like politicians, and they’ll tell you they like politicians too. But the difference in attitudes elicited isn’t a consequence of any change in attitude toward any fixed group of people. The bystander information causes a change in the perceived referent of “politicians” (from local party hacks to great statespersons).

Just as with the literature on the effects of bystander information on attitude, most of the literature in social psychology on conformity has taken the differences in attitudes across conditions to be a manifestation of changed attitudes to a common object of judgment. There is only a small cluster of papers (published mostly in the 1980s and 1990s), which expressly investigate the prospects of the change-in-the-object-of-judgment explanation (namely: (Allen & Wilder, 1980), (Griffin & Buehler, 1993), (Buehler & Griffin, 1994), (Wood, Pool, & Purvis, 1996), (Pool, Wood, & Leck, 1998), (Cohen, 2003), (Hayes, Lee, & Wood, 2018)). Their findings are tantalizing.

In Allen and Wilder's (1980), participants were asked if they agreed with the statement "I would never go out of my way to help someone in need." But in one condition they were told most people agreed with this statement and in another condition they were told that most people disagreed with it. Sure enough, apparent conformity of attitude was witnessed. But so too was a change in the interpretation of the statement (specifically "go out of my way"): between putting oneself in harm's way (when bystanders agreed with it) to engaging in some minor inconvenience (when bystanders disagreed with it).

Buehler and Griffin ((1993), (1994)) were interested in why bystander information changes interpretation. They distinguished between two possibilities: (i) people are forming an attitude toward the stimulus and then changing their interpretation of the stimulus in order to "justify" the relation of their attitude to the majority attitude (conformity or deviance); (ii) people are using bystander information as evidence to help them interpret the stimulus, and then forming their attitude toward the stimulus so interpreted. Bystander information can function as evidence of the meaning of the stimulus most likely in multiple ways. But one way is: if you know already what bystanders believe (and so would agree with), then you can narrow down how they are interpreting the stimulus by checking what interpretations are entailed by what they believe.

Buehler and Griffin's most commonly used stimulus was something called "Robert's Dilemma": Robert has the chance to study music instead of continuing along a stable career in medicine. After reading about Robert's predicament, participants are asked what the lowest chance of success is when it would still be worth Robert trying the music career (their chance-attitude). Bystander information is information about what others think of this chance of success. Buehler and Griffin tried out different sequences of elicitation:

- first they provided bystander information, then they elicited the participant's interpretation of the description of Robert's predicament, then they elicited the participant's chance-attitude.
- first they provided bystander information, then they elicited the participant's chance-attitude, then they elicited the participant's interpretation of the description of Robert's predicament.

Buehler and Griffin found no effect of bystander information on how participants interpreted the description of Robert's predicament unless the interpretation was elicited *after* the participants' own chance-attitudes were expressed. In this condition, when told a majority had a "risky" view (that even with a low chance of success, Robert should go for it), those who agreed with the bystanders interpreted the description of Robert's predicament as making it seem more reasonable for him to go for the music career despite the risk of failure: e.g. because Robert valued the things he'd get from a music career far more highly than the things he'd get from a medical career. Those who disagreed with the bystanders interpreted the description of Robert's situation as making it seem more reasonable for him to stick to his medical training: e.g. because Robert valued the things he'd get from the medical career more than he'd value the things he'd get from the music career. But as I said: this result was found only for participants whose interpretations were elicited after they had expressed their chance-attitudes. For this reason, Buehler and Griffin conclude that their participants were interpreting in order to make their position relative to the bystanders seem more reasonable (rather than because bystander information constituted evidence of what the object of judgment is).

Wood et al. ((1996), (1998)) argue that Buehler and Griffin are wrong to conclude that bystander information shapes interpretation of a linguistic stimulus only if the interpreter has already publicly expressed their own attitude toward the stimulus. What matters, they believe, is whether participants have sufficient opportunity and motivation to interpret in light of bystander information. If given such opportunity and motivation before they express their own attitude, then bystander information will shape interpretation before expression of attitude. This is indeed what Wood et al. found. For instance, one stimulus was: “Sex of employees should be considered in job promotion decisions.” If told that a group called the Radical Lesbian Feminists strongly disagreed with this statement, then those who felt strongly that they didn’t want to be associated with this group interpreted the statement as permitting consideration of sex-related characteristics, if that’s relevant to the job, and then agreed with the statement. Those who didn’t care whether they were similar to the Radical Lesbian Feminists were less likely to adopt such an interpretation: instead interpreting the statement as meaning that under no circumstances, and in no way, should sex be considered.

The researchers carrying out the studies summarized in section 3 observed differences in attitudes toward a common post when that post is accompanied by different bystander information. They conclude from this that bystander information *changes* attitudes. But, as we can now see, bystander information has the capacity to change what we think the content of a linguistic stimulus (such as a post) is. This problematizes the inference of researchers whose work is summarized in section 3 just as it problematizes the same inference in conformity studies in social psychology more generally. It makes live the possibility that bystander information changes what we take the content of a post to be rather than changing attitudes towards any constant content.

## **5 Three Mechanisms of Bystander Influence on Linguistic Interpretation**

The studies summarized in section 4 provide evidence that bystander information can shape linguistic interpretation. But each cluster of studies is suggestive of a different mechanism by means of which bystander information can do so. In this section, I describe three mechanisms suggested by these studies: “evidence-based interpretation”; “social-placing interpretation”; and “doubling-down interpretation”.

### **5.1 Evidence-Based Interpretation**

Firstly, suppose you know that a sentence *S* could be used to express two propositions: *A* and *B*. You know that bystanders believe *A* to be true but *B* to be false. You then learn that these bystanders disagree with the sentence *S*. From this you can infer that the bystanders are not interpreting *S* to mean *A*: they must understand *S* to mean *B*. Thus knowledge of what bystanders believe, what a sentence could be understood to mean (a range of propositions) and what attitude the bystanders have toward the sentence, can together be used to figure out how the sentence is being interpreted by the bystanders. You can then use this as the basis for establishing what object (content) you are being asked to express your attitude about.<sup>7</sup> This kind of thing seems to be the mechanism behind the results of

---

<sup>7</sup> This resembles cases in which you see for yourself what’s being described, and infer from this conclusions about with which contents constructions in the description are being employed (Barker, 2002): you see how tall Anja is

Allen and Wilder (1980): where the construction “go out of my way” was interpreted differently depending upon whether participants were told that an anonymous majority agreed or disagreed with the sentence “I would never go out of my way to help someone in need.” The additional factors introduced by subsequent studies in order to incentivize reinterpretation (namely: bystander identity and the relation of one’s attitude to a majority attitude) are absent in this case. Yet a difference of interpretation across conditions was found. The simplest explanation is that bystander information was being used to triangulate the content of the linguistic stimulus given what’s known about the bystanders: it’s because people assume that most people think that they *would* inconvenience themselves to help someone in need, but do *not* think that they would risk their lives to help someone in need, that we find one interpretation of the sentence “I would never go out of my way to help someone in need” when told bystanders agree with it, and another interpretation when bystanders disagree with it.

## 5.2 Social-placing Interpretation

Secondly, linguistic interpretation may be driven by a desire to express the same attitude (agreement/disagreement) toward a sentence as those who you want to be like express towards it, and to express a different attitude (agreement/disagreement) toward a sentence than that which those who you don’t want to be like express towards it. Suppose a sentence can bear multiple contents: A or B. You agree with A and you disagree with B. You then learn that people you want to be like agree with the sentence. You may then be motivated to interpret the sentence as expressing A, so that you can express the same attitude toward the sentence as those who you want to be like. Had you learned instead that people you don’t want to be like agreed with the sentence (or had you learnt that people you want to be like disagree with the sentence), then you would have interpreted it as expressing B instead: so that you could locate yourself where you want to be relative to these groups *vis a vis* whether you agree/disagree with this sentence.

This is how Wood et al. ( (1996), (1998)) understand their own findings concerning sentences like “Sex of employees should be considered in job promotion decisions”. Wood et al. made the identity of bystanders salient. The bystanders were made to be people from whom one would want to distance oneself (or not); e.g. Radical Lesbian Feminists. And those who felt most strongly about the identity of these people were the people who changed construal prior to expression of attitude (e.g. in whether the statement prohibits appeal to sex-related characteristics, if they’re relevant to the job). Wood et al. concluded that their independent variable (identity of bystanders), given a background factor (the desire to be similar to/different from the identity of the bystanders), causes the change of interpretation.<sup>8</sup>

---

when you’re told “Anja’s tall” and from this you infer something about the content of “tall” in this context. In the current case, instead of using your own direct knowledge of what’s being described to draw conclusions about with which contents constructions in the description are being employed, you use your knowledge of what others think in order to reach the same sort of conclusion.

<sup>8</sup> I grant Wood et al.’s own interpretation of their data in the main body of this paper. But I want to register here that their results don’t fully support this interpretation: any other variable that may be correlated with how one feels about the bystanders may well explain changes of interpretation in light of this bystander information. Amongst such correlates are your beliefs about what bystanders (with the identity you are told they have) already believe. For this reason, Wood et al.’s experiment design doesn’t rule out the possibility that participants are



### 5.3 Doubling-down Interpretation

Thirdly, there's doubling-down interpretation. With social-placing interpretation, one looks for an interpretation that allows one to agree (disagree) with those one wants to agree (disagree) with before one has expressed one's own attitude. With doubling-down interpretation, one has already expressed one's attitude toward a sentence and one already knows what others' attitudes are towards it. But one looks for an interpretation that makes the relation of one's attitude to that of others (conformity or deviance) seem more reasonable. Think for instance of saying something in a seminar and prominent people in the room then disagree with it. One falls into a pit of social anxiety. To make oneself feel better, one rehearses an interpretation of what was at issue so that one's position seems (more) reasonable: paying little heed to whether one shares content with others (preoccupied as one is with calming oneself down). Buehler and Griffin (1993), (1994) take themselves to be witnessing this mechanism in action when they witness the influence of bystander information on participants' interpretations of the description of Robert's situation when he faces the dilemma about whether to pursue a music career: those who feel most anxious about disagreeing with others adapt their interpretation of Robert's situation to make their own position seem more reasonable.<sup>9</sup>

## 6 Situating the Mechanisms in Social Media

We have in view three mechanisms by means of which linguistic interpretation can be shaped by bystander information. I now want to work through how these modes of interpretation could jeopardize the sharing of content on social media.

### 6.1 Evidence-based Interpretation in Social Media

Let's start with the practice of evidence-based interpretation. Here, bystander information is used for the evidential value it has in order to better understand what's being said: a reversal of what we might

---

engaging in evidence-based interpretation. In this respect, Wood et al.'s inference is typical of research into motivated cognition: which typically doesn't control for doxastic confounds that can explain differential responses to the same evidence (for discussion see Begby (2021), Borg (2022), Davies (forthcoming), and Tappin, Pennycook, and Rand (2020a)).

<sup>9</sup> What was said in the previous footnote about Wood et al.'s studies can also be said about most of Buehler and Griffin's: they don't demonstrate that evidence-based interpretation isn't in action. However, there is one study of theirs (study 4 in (Griffin & Buehler, 1993)), which provides far stronger evidence of a doubling-down interpretation. In this study, whether or not participants agreed with the bystanders was experimentally manipulated by asking participants to give their judgments first (on Robert's dilemma), and then they were given bystander information which differed by 20% either upwards (that the minimal chance of success that would make it worth Robert going for the music career was 20% higher than the participant thought) or downwards (that the minimal chance of success was 20% lower than the participant thought). Sure enough, those who were riskier than the bystanders adopted an interpretation of Robert's situation that justified taking more risk and those who were less risky than the bystanders adopted an interpretation of Robert's situation that justified taking less risk.

otherwise have thought of as the natural order of things—wherein you figure out what someone says in order to figure out who agrees and who disagrees with it. But it is nonetheless something one can do deliberately without feeling guilt for one’s epistemic conduct: in contrast with social-placing interpretation and doubling-down interpretation, evidence-based interpretation is a matter of following the evidence where it leads. So how could it threaten content-sharing? I describe two ways in which evidence-based interpretation can lead to failures to share content.

First: when there’s misleading evidence. What kind of misleading evidence would scupper unitary interpretation? One answer: evidence that generates false beliefs about what bystanders believe. Suppose a Democrat reasonably, but falsely, believes that Republicans are against migration in any form. A post with the sentence, “Better migration controls proposed” yields a lot of agreement from a Republican audience—the bystanders. Given the Democrat’s beliefs about the beliefs of Republicans, and given that Republicans are agreeing with this post, the Democrat will be likely to conclude that “Better migration controls” is being understood to bear a content like: controls which reduce immigration as much as possible, regardless of what other consequences such a move would entail. But the Democrat’s belief about Republicans’ beliefs is false. Republicans are typically of the view that some migration has positive benefits and shouldn’t be eliminated altogether (Yudkin, Hawkins, & Dixon, 2019). Thus, Republicans and Democrats will develop divergent assignments of content to this sentence in part because the presentation of bystander information alongside the post unlocks the influence of a Democrat’s false beliefs about Republicans’ beliefs—allowing their influence to ripple out into the Democrat’s interpretation of the post. If the bystander information had not been present, then the false beliefs the Democrat has about Republicans would not have influenced their interpretation of the post. In this sense, bystander information unlocks the bad hermeneutical influence of false beliefs about the beliefs of others (e.g. political opponents). That influence would not have been had if the bystander information had not been present. False polarization breeds additional false polarization.

Secondly, we can distinguish between a post type and a post occurrence: e.g. an identical quotation (post type) might appear on different people’s Facebook walls (different occurrences). Evidence-based interpretation implies that there’ll be a path-dependence to how each occurrence of the same post is interpreted. Suppose that the first responders to each occurrence of the same quotation (“Doctors should go out of their way to help someone in need”) diverge in their attitudes towards it (one agrees, the other disagrees). They express this: thereby adding bystander information to each occurrence of the same quotation. When new readers come along, if they engage in evidence-based interpretation, they’ll use the existing bystander information to help them interpret the occurrence of the quotation. But since the bystander information differs, the interpretation that evidence-based interpretation will lead them toward will differ. Consequently, even if readers of each occurrence of the quotation have the same attitudes towards the same contents, they may well express divergent attitudes towards this same quotation (because they differ in their interpretation of its content). They express their attitudes and the process repeats. Thus, evidence-based interpretation implies a path-dependence of responses to a post. If a single post (a headline, a quote, etc.) occurs in two isolated contexts on social media, given an element of randomness in how first responders respond, there’ll not uncommonly develop a different set of bystander judgments for each occurrence of the post. Given the evidence-based mechanism, this could easily lead subsequent readers of the post to divergent content assignments for the post. Again: if there were no bystander information, we wouldn’t expect the same. What enables this to happen is

both: that a single post can appear in multiple contexts on social media and that in each context it is accompanied by bystander information.<sup>10</sup>

## 6.2 Social-placing Interpretation on Social Media

Imagine someone who deliberately seeks out information from sources from “across the aisle.” One way in which this attempt to balance their informational diet can fail is that they simply have lower credence in information coming from such sources (aka “echo chambers” in Nguyen’s (2020) sense). That’s what we might call an *epistemic* mechanism for muting the impact of this information—a mechanism that thwarts one’s attempts to broaden one’s informational diet.

But an alternative *hermeneutic* mechanism takes the form of social-placing interpretation. Imagine again someone who deliberately seeks out information from sources from “across the aisle.” Though a Democrat, she follows a source on social media from the other side of the political spectrum. There she sees a post: “Study finds that the Texan economy is dependent on Mexico.” We can distinguish between two kinds of dependence which this sentence could be claiming to exist: the Texan economy depends on the work of illegal Mexican migrants, or, the Texan economy depends on trade agreements between the US and Mexico. Suppose that the Democrat believes that the Texan economy’s dependence on illegal migrant labour is so great that the benefits Texas receives from any trade agreement are negligible. Suppose that she sees that huge numbers of Republicans agree with the sentence. Not wanting to affiliate with them, she is drawn to an interpretation of the sentence she can disagree with: that it’s asserting a dependence via existing trade agreements. If she interprets in this way, her interpretation is not reached by any attempt to understand what the Republicans are actually agreeing with. For that reason, it would have to be a coincidence if it turned out that what she was disagreeing with is what the Republicans were agreeing with. She’s not driven by the goal of understanding what was said. She’s driven by the desire to place herself socially relative to these Republicans. Thus, the risk to content-sharing is high.

## 6.3 Doubling-down Interpretation on Social Media

On social media, we read and—without any great effort—express to other readers our attitudes toward the posts we read. This feature of reading on social media can in principle incentivize one to engage in doubling-down interpretation. For instance, if you react to a post (with a reaction or a comment), and subsequently discover that most (respected) others take a different attitude toward the post, you will experience pressure to interpret the post so that your deviant reaction seems more reasonable.

Suppose you see a post about a late philosopher, pleading with others to acknowledge his reported misconduct when he was alive. You like the post—thinking that it’s good for such things to be acknowledged and that someone with authority manages to say such a thing. But then you see that

---

<sup>10</sup> I’m not denying that the same effect could arise out of attitudinal conformity without reinterpretation of the post. What I’m saying is that, once we acknowledge the effect that bystander information can have on interpretation (and, in this particular case, via what I’m calling “evidence-based interpretation”), we should take seriously the possibility that bystander information can undermine the sharing of content on social media (even when the vehicle of content is successfully shared).

others vehemently disagree with this post, putting forward various reasons for doing so in the comments: including that his behaviour was reportedly a by-product of his ASD. This makes you feel social anxiety. This may cause you to rescind your like as you consider the issue further. But it may instead lead you to keep the like and seek out and rehearse to yourself an interpretation of the original post on which your like—despite the disagreement by others—seems more reasonable. Suppose for instance that the original post includes the sentence “he drove women out of the profession.” We can distinguish two ways to understand this sentence. On one, it is true only if he culpably engaged in conduct which caused women to leave the profession. On the other, it is true even if the conduct which caused women to leave the profession was not culpable (i.e. the second interpretation is weaker than the first). One might switch one’s interpretation from the stronger to the weaker interpretation when met with disagreeing bystanders in order to make oneself feel that one’s agreement (like) of the post is more reasonable—that what one meant to agree with was not all that much. You interpret in order to relieve yourself of the discomforts of social anxiety, losing sight of whether this is the object of others’ judgment. If one arrives at one’s interpretation in this way, there’s a substantial risk of failing to share content—since this is clearly not a reliable way of grasping what was meant by the author or what was understood by the bystanders.

#### **6.4 Risk factors**

The displacement of content-sharing by ignorance-affected evidence-based interpretation, social-placing interpretation and doubling-down interpretation depends upon the presence of other factors. We can distinguish between two groups of risk factor.

Firstly, there are those described in this section: false beliefs about bystanders’ beliefs; desires to place oneself socially; and a proclivity to experience social anxiety when discovering that one’s attitudes towards given posts don’t align with those of bystanders who one wishes to be like. For those who don’t have these traits, the modes of interpretation just canvassed are less likely to be invoked and therefore less likely to distract from the business of content-sharing.

Secondly, think back also to the mediating factors listed in section 3. Those were factors that were preconditions on bystander information’s influence on attitudes towards a post. True: those factors were understood, by the experimenters uncovering them, as preconditions on the capacity of bystander information to change social media users’ minds about a fixed object of judgment. But if those experimenters are (to some extent at least) jumping the gun—misinterpreting a change in the object of judgment as a change of mind—then these factors may well be preconditions on the capacity of bystander information to change social media users’ interpretation of the content of a post: in which case, these too will be risk factors *vis a vis* the derailing of content-sharing on social media.

#### **7 Social Media’s *Laissez Faire* Socialization of Reading**

It should now be clear how bystander information on social media could serve to hinder content-sharing. There’s empirical evidence that bystander information shapes linguistic interpretation. That evidence suggests a variety of interpretative mechanisms by means of which this happens. Each mechanism either constitutes a form of interpretation that isn’t at all directed at grasping a shared

content (as with social-placing and doubling-down interpretation) or it is, but it can malfunction, either because of the presence of false beliefs about what others think, or because of the structure of social media (as with evidence-based interpretation).

However, the twin facts that social media forces bystander information upon us and that bystander information can trigger these modes of linguistic interpretation are not sufficient for us to conclude that social media constitutes a kind of context that hinders content-sharing. If social media threatens content-sharing (in the way that we're exploring), that's going to be because it combines bystander information with something else. In this final section, I describe what that something else is.

The socialization of reading is not unique to social media: the quintessential solitary reader is merely quintessential. There are many contexts for reading wherein you're not alone in your attitudes towards what's read. To name a few, there are: village book clubs; university courses; giggling friends around an unfortunately worded instant message; and any team-based production of a text (an advertising team, a team of journalists and their editors, those producing documents for internal use within a company etc.). In each such context, your reading of a text is accompanied by bystander information and your judgments about it are publicised to those present in the context.

However, not all contexts that socialize reading are alike in how they do so: some contexts socialize reading in a way which (by and large) washes out the splintering influence of bystander information on interpretation across different readers. Others don't.

Compare the reading of a text on social media with the reading of a text as part of a university course. Think for instance of a course on Wittgenstein's *Tractatus Logico-Philosophicus*. In most courses on this text, students will be penalized in ways they care about if they deviate from rather specific sets of interpretations. The course is, in part, a training in what contents it is permissible to ascribe to (particular passages of) this text—in how to read it. This training is enacted through social sanctions and commentary, encouragement and criticism from someone with the power to decide a grade, upon which, some portion of a student's self-esteem and future depends. The prior investments students have made in order to be there (in time, stress and finance) help filter course attendees down to those with an interest in getting appropriate accreditation, which can be had only by successfully undergoing the training the course offers. Because of the professional pressures they are under, those running and grading such courses (across multiple institutions worldwide) won't make arbitrary sets of contents permissible for ascription to the *Tractatus's* key passages. The readings they permit will be those thought permissible by the broader professional philosophical community. While social-placing, doubling-down and ignorance-affected evidence-based interpretation no doubt are operative within such university courses, succumbing to their influence won't help the student get what he (presumably) wants. A student who doubles down on an interpretation of the *Tractatus* because he feels anxiety in the face of his classmates or adopts an interpretation to allow him to distance himself from, or side himself with, others in the classroom, won't, for all that, raise his grade. He knows that. Which is why he's less likely to do it.

The way in which social media socializes the reading of a text is broadly *laissez faire* by comparison: the environment that serves (in principle) to wash out the influence of social-placing, doubling-down and ignorance-affected evidence-based interpretation in a university course is, in many places, absent. By and large, social media doesn't have a structure that unifies—either through stick-and-carrot incentives or by filtering who's allowed entry—the purposes and background knowledge (in short: the common-

ground) of those participating in the exchanges thereon. Yet it does introduce bystander information: something with the potential to drive apart the purposes of those interpreting one and the same post.

Of course there are exceptions. There are successful (closed) professional philosophical Facebook groups wherein the factors that help to homogenize interpretation of relevant texts in professional philosophical circles extend to social media (the importance of professional appraisal, the extreme filtering that determines who knows about and has reason to be a member of these groups). And on the other side of things—of course there are going to be cases when the professional and organizational pressures that can wash out the influence of identity, anxiety and ignorance, fail to do so.

But none of this slights the fact that in large regions of social media, the pressures that make indulgence of social-placing, doubling-down and ignorance dangerous to one's own interests in (for instance) the university course context, are absent. If social media does hinder content-sharing in the way articulated in this paper, it does so because it injects bystander information into the activity of reading without introducing such antidotal pressures.

## References

- Abreu Zavaleta, M. (2019). Communication and Variance. *Topoi*.
- Allen, V. L., & Wilder, D. A. (1980). Impact of Group Consensus and Social Support on Stimulus Meaning: Mediation of Conformity by Cognitive Restructuring. *Journal of Personality and Social Psychology*, 39(6), 1116-1124.
- Asch, S. (1940). Studies in the Principles of Judgments and Attitudes: II. Determination of Judgment by Group and by Ego Standards. *Journal of Social Psychology*, 12, 433-465.
- Asch, S. (1948). The Doctrine of Suggestion, Prestige and Imitation in Social Psychology. *Psychological review*, 55, 250-276.
- Asch, S. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetskow (Ed.), *Groups, Leadership and men; research in human relations* (pp. 177-190). Carnegie Press.
- Barker, C. (2002). The Dynamics of Vagueness. *Linguistics and Philosophy*, 25, 1-36.
- Begby, E. (2021). *Prejudice: A Study in Non-Ideal Epistemology*. Oxford: Oxford University Press.
- Bond, R. (2005). Group Size and Conformity. *Group Processes and Intergroup Relations*, 8(4), 331-354.
- Bond, R., & Smith, P. B. (1996). Culture and Conformity: A Meta-Analysis of Studies Using Asch's (1952b, 1956) Line Judgment Task. *Psychological Bulletin*, 119(1), 111-137.
- Borg, E. (2022). In Defence of Individual Rationality. *Proceedings of the Aristotelian Society*, 122(3).
- Bowker, M. (2019). Ineliminable underdetermination and context-shifting arguments.
- Buehler, R., & Griffin, D. W. (1994). Change-of-Meaning Effects in Conformity and Dissent: Observing Construal Processes Over Time. *The Journal of Personality and Social Psychology*, 67(6), 984-996.

- Chung, M. (2017). Not just numbers: The role of social media metrics in online news evaluations. *Computers in Human Behavior, 75*, 949-957.
- Chung, M. (2019). The message influences me more than others: How and why social media metrics affect first person perception and behavioral intentions. *Computers in Human Behavior, 91*, 271-278.
- Cohen, G. (2003). Party of Policy: The Dominating Impact of Group Influence on Political Beliefs. *The Journal of Personality and Social Psychology, 85*(5), 808-822.
- Crutchfield, R. S. (1955). Conformity and Character. *American Psychologist, 191*-198.
- Davies, A. (2019). Testimonial Knowledge and Context-Sensitivity: a New Diagnosis of the Threat. *Acta Analytica, 34*(1), 53-69.
- Davies, A. (2021). Communicating in contextual ignorance. *Synthese, 199*, 12385-12405.
- Davies, A. (forthcoming). Science Communication, Cultural Cognition and the Pull of Epistemic Paternalism. *Journal of Applied Philosophy*.
- Dixon, G. (2020). Undermining Credibility: The Limited Influence of Online Comments to Vaccine-Related News Stories. *Journal of Health Communication, 25*(12), 943-950.
- Frost-Arnold, K. (2021). The Epistemic Dangers of Context Collapse Online. In J. Lackey (Ed.), *Applied Epistemology* (pp. 437-456). Oxford: Oxford University Press.
- Griffin, D. W., & Buehler, R. (1993). Role of Construal Processes in Conformity and Dissent. *The Journal of Personality and Social Psychology, 65*(4), 657-669.
- Grylls, G. (2022, April 12). Estonia asks for more British troops to deter possible Russian attack. *The Times*. Retrieved 12 31, 2022, from <https://www.thetimes.co.uk/article/estonia-asks-for-more-british-troops-to-deter-possible-russian-attack-dqgz208mm>
- Hayes, T., Lee, J. C., & Wood, W. (2018). Ideological group influence: central role of message meaning. *Social Influence, 13*(1), 1-17.
- Hinnant, A., Subramanian, R., & Young, R. (2016). User comments on climate stories: impacts of anecdotal vs. scientific evidence. *Climate Change, 138*, 411-424.
- Kim, H., Han, J., & Seo, Y. (2020). Effects of Facebook Comments on Attitude Toward Vaccines: The Roles of Perceived Distributions of Public Opinion and Perceived Vaccine Efficacy. *Journal of Health Communication, 1*(25), 159-169.
- Kim, Y. (2015). Exploring the Effects of Source Credibility and Others' Comments on Online News Evaluation. *Electronic News, 9*(3), 160-176.
- Kluck, J., Schaewitz, L., & Krämer, N. (2019). Doubters are more convincing than advocates. The impact of user comments and ratings on credibility perceptions of false news stories on social media. *Studies in Communication and Media, 8*(4), 446-470.

- Lee, E., & Jang, Y. (2010). What do others' reactions to news on Internet portal sites tell us? Effects of presentation format and readers' need for cognition on reality perception. *Communication Research, 37*(6), 825-846.
- Lee, E.-J., Jang, Y. J., & Chung, M. (2021). When and How User Comments Affect News Readers' Personal Opinion: Perceived Public Opinion and Perceived News Position as Mediators. *Digital Journalism, 9*(1), 42-63.
- Lee, J. (2014). Are some people less influenced by others' opinions? The role of internal political self-efficacy and need for cognition in impression formation on social networking sites. *Cyberpsychology, Behavior, and Social Networking, 17*(9), 571-577.
- Lee, S., Atkinson, L., & Sung, Y. (2022). Online Bandwagon Effects: Quantitative versus qualitative cues in online comments sections. *New Media & Society, 24*(3), 580-599.
- Longworth, G. (2018). Understanding what was said. *Synthese, 195*(2), 815-834.
- Nguyen, C. T. (2020). Echo Chambers and epistemic bubbles. *Episteme, 17*(2), 141-161.
- Nguyen, C. T. (2021). How Twitter Gamifies Communication. In *Applied Epistemology*. Oxford: Oxford University Press.
- Pariser, E. (2011). *The Filter Bubble: How the new personalized web is changing what we read and how we think*. London: Penguin Books.
- Peet, A. (2016). Testimony and the epistemic uncertainty of interpretation. *Philosophical Studies, 173*(2), 395-416.
- Pool, G., Wood, W., & Leck, K. (1998). The Self-Esteem Motive in Social Influence: Agreement with Valued Majorities and Disagreement With Derogated Minorities. *The Journal of Personality and Social Psychology, 75*(4), 967-975.
- Porten-Che e, P., & Christiane, E. (2020). The effects of likes on public opinion perception and personal opinion. *Communications, 45*(2), 223-239.
- Record, I., & Miller, B. (2022). People, posts, and platforms: reducing the spread of online toxicity by contextualizing content and setting norms. *Asian Journal of Philosophy, 1*, 1-19.
- Sah, Y., & Peng, W. (2022). Context-dependent Online Social Influence: Effect of Majority and Minority Comments on Posters and Lurkers. *Social Science Computer Review*.
- Simon, H. (1954). Bandwagon and underdog effects and the possibility of election predictions. *Public Opinion Quarterly, 18*(3), 245-253.
- Sundar, S. S. (2008). The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility. In M. J. Metzger, & A. J. Flangagin (Eds.), *Digital Media, Youth, and Credibility* (pp. 73-100). Cambridge, MA: MIT Press.
- Tappin, B., Pennycook, G., & Rand, D. (2020a). Thinking clearly about causal inferences of politically motivated reasoning: why paradigmatic study designs often undermine causal inference. *Current Opinion in Behavioral Sciences, 34*, 81-87.



- Wijenayake, S., Hettiachchi, D., Hosio, S., Kostakos, V., & Goncalves, J. (2021). Effect of Conformity on Perceived Trustworthiness of News in Social Media. *IEEE Internet Computing, 25*(1), 12-19.
- Winter, S., Brückner, C., & Krämer, N. C. (2015). They Came, They Liked, They Commented: Social Influence on Facebook News Channels. *Cyberpsychology, Behavior, and Social Networking, 18*(8), 431-436.
- Winter, S., Krämer, N., & Liang, Y. (2017). User-generated opinion: How reader reactions and source reputation influence the effects of online news. *Studies in Communication and Media, 6*(3), 240-261.
- Wood, W., Pool, G. L., & Purvis, D. (1996). Self-Definition, Defensive Processing, and Influence: The Normative Impact of Majority and Minority Groups. *The Journal of Personality and Social Psychology, 71*(6), 1181-1193.
- Yudkin, D., Hawkins, S., & Dixon, T. (2019). *The Perception Gap: How False Impressions Are Pulling Americans Apart*. New York: More in Common.
- Zerback, T., & Töpfl, F. (2022). Forged Examples as Disinformation: The Biasing Effects of Political Astroturfing Comments on Public Opinion Perceptions and How to Prevent Them. *Political Psychology, 43*(3), 399-418.