The Role of Trust in Argumentation

CATARINA DUTILH NOVAES

Philosophy
Vrije Universiteit Amsterdam
1081 HV Amsterdam
The Netherlands
c.dutilhnovaes@vu.nl

Arché
Philosophical Research Centre for Logic,
Language, Metaphysics and Epistemology
The University of St Andrews
17–19 College Street, St Andrews, Fife
Scotland
KY16 9AL

Abstract: Argumentation is important for sharing knowledge and information. Given that the receiver of an argument purportedly engages first and foremost with its content, one might expect trust to play a negligible epistemic role, as opposed to its crucial role in testimony. I argue on the contrary that trust plays a fundamental role in argumentative engagement. I present a realistic social epistemological account of argumentation inspired by *social exchange theory*. Here, argumentation is a form of *epistemic exchange*. I illustrate my argument with two real-life examples: vaccination hesitancy, and the undermining of the credibility of traditional sources of information by authoritarian politicians.

Résumé: L'argumentation est importante pour le partage des connaissances et des informations. Étant donné que le destinataire d'un argument s'engage prétendument d'abord et avant tout dans son contenu, on peut s'attendre à ce que la confiance joue un rôle épistémique négligeable, par opposition à son rôle crucial dans un témoignage. Je soutiens au contraire que la confiance joue un rôle fondamental dans l'engagement argumentatif. Je présente une description épistémologique sociale réaliste de l'argumentation inspirée de la théorie des échanges sociaux. Ici, l'argumentation est une forme d'échange épistémique. J'illustre mon argumentation par deux exemples concrets: l'hésitation à la vaccination et l'atteinte à la crédibilité des sources d'information traditionnelles par les politiciens autoritaires.

Keywords: trust, argumentation, testimony, social exchange theory, epistemic vigilance

1. Introduction

Argumentation, understood as the practice of "giving and asking for reasons" (Brandom, 1994), seems to be an important mechanism for sharing knowledge and information. When being presented with an argument, a knower can come to know something new—the conclusion—if she believes the premises and recognizes the relation of epistemic support between premises and conclusion. A paradigmatic case of transfer of knowledge by means of argumentation occurs in mathematics, where mathematical proofs are quintessential instruments for sharing mathematical knowledge (Easwaran 2009) (Dutilh Novaes 2021, Chapter 11). But the phenomenon is a more general one, and indeed argumentation seems to play a pivotal epistemic role in a wide range of domains such as science, politics, and law, among others.

Argumentation can be contrasted with testimony in that, through testimony, I may acquire knowledge solely by aptly judging the testifier to be a reliable informant without inspecting her own grounds for believing some proposition. With argumentation, however, my estimation of my informant's reliability is not (purportedly at least) what grounds the transfer of knowledge. Instead, it is my own endorsement of the premises and of the relation of epistemic support between premises and conclusion that justifies my newly acquired belief in the conclusion. Thus understood, one would expect trust to play a negligible epistemic role in argumentation as opposed to testimony where it does play a crucial role.

However, one may wonder whether there really is such a difference between testimony and argumentation with respect to trust; maybe trust is also a fundamental element in argumentative processes (descriptively and/or normatively speaking). For example, what should we say about the receiver of an argument who dismisses it outright, with no real engagement with its content, simply because he does not trust the source? Is he breaking some epistemic norm? Can this response be justified at all? In order to address these questions, a promising avenue is to approach argumentation from the perspective of social epistemology.

Perhaps surprisingly, there has been relatively little attention paid to argumentation as a topic of investigation in (analytic) social

epistemology. The notable exception is A. Goldman, who has written extensively on the topic (Goldman 1994; 1997; 2004). However, while foundational and groundbreaking, Goldman's account of the social epistemology of argumentation is in fact highly idealized, and not aimed at dealing with argumentation in "messy" real-life scenarios. In particular, it is arguably not fully suitable to address the role of trust in argumentation because, among other reasons, it seems to postulate that we readily engage in argumentation with interlocutors who fundamentally disagree with us. But this is far from certain, both descriptively and normatively. What we need instead is a more realistic social epistemology of argumentation to investigate argumentation in real-life scenarios.

In this paper, I formulate such a theory inspired by the framework of *social exchange theory* as developed by sociologists and social psychologists since the mid-20th century (Cook 2013). The key idea is that argumentation can be seen as a form of *epistemic exchange*, and so the findings from social exchange theory concerning exchanges in general should also be relevant for a special kind of exchange, namely epistemic exchange, and in particular for argumentation as one modality (among others) of epistemic exchange. Specifically, social exchange theorists maintain that trust is a fundamental component of exchange relations in general (both descriptively and normatively speaking). If this is true of exchanges in general, it should also hold for argumentation as a kind of (epistemic) exchange. Thus, on the basis of a conceptualization of argumentation as epistemic exchange, I argue that trust does indeed play a more significant role in argumentation than is usually recognized in the literature: arguers will tend to be wary of engaging in argumentation with interlocutors whom they do not trust sufficiently. Importantly, I discuss both descriptive and normative aspects of these questions, but the goal is ultimately to combine the two perspectives so as to arrive at a *prescriptive* account of the social epistemology of argumentation.[1]

---

[1] A prescriptive account consists in "how we might help less rational people, who nevertheless aspire to rationality, to do better" (Bell, Raiffa, and Tversky 1988, p. 9). While the concept of rationality may be contentious, the idea is simply to help imperfect agents to "improve" (on some relevant metric) their behavior in a

The paper proceeds as follows: I start by discussing the concept of epistemic vigilance and locating testimony and argumentation within a range of different strategies for epistemic vigilance. I then turn to argumentation specifically and argue that a Millian conception of argumentation, according to which engaging with dissenters will naturally lead to epistemic improvement, is descriptively inaccurate. In section 4, I briefly present social exchange theory, and in section 5, I sketch a conception of argumentation as a form of epistemic exchange. I then, in section 6, discuss whether argumentation can indeed compensate for diminished trust in situations of epistemic exchange, as argued by some authors, and conclude that it cannot. I also introduce a three-tiered model for argumentative engagement. In section 7, I discuss some concrete examples where argumentation and trust (or the lack thereof) are closely related.

2. Epistemic vigilance, testimony, and argumentation

Humans are hyper-social animals. We need each other to survive as opposed to many animal species that lead solitary lives. Moreover, our epistemic lives are also fundamentally social; most of what we know, we learn from others (Harris 2012). However, the social nature of our epistemic lives entails vulnerability: information shared by others may not always be accurate, be it because informants themselves are mistaken in their beliefs or because they deliberately misinform us to their own benefit, so we must also have ways to assess the quality of the information conveyed to us by others.

Inspired by these observations, Sperber and colleagues have introduced the useful concept of *epistemic vigilance*:

> Humans massively depend on communication with others, but this leaves them open to the risk of being accidentally or intentionally misinformed. To ensure that, despite this risk, communication remains advantageous, humans have, we claim, a suite of cognitive mechanisms for epistemic vigilance (Sperber, Clément, Heintz, Mascaro, Mercier, and Origgi 2010, p. 359).

certain domain so that the gap between the descriptive and the normative becomes smaller.

Epistemic vigilance is an umbrella term that is meant to cover a wide range of different cognitive mechanisms/strategies. These include, for example, ways of tracking and assessing people's trustworthiness as sources of information by means of reputational elements, the ability to infer intentions in others, and coherence checks of the message content against the background of prior beliefs. Sperber et al. believe that we are well equipped to perform these cognitive tasks, thus being by and large competent when it comes to epistemic vigilance. For our purposes, one distinction drawn by Sperber et al. is particularly significant:

> [E]pistemic vigilance can be directed at the *source* of communicated information: Is the communicator competent and honest? It can also be directed at the *content* of communication, which may be more or less believable independently of its source (Sperber, Clément, Heintz, Mascaro, Mercier, and Origgi 2010, p. 375, emphasis added).

Consider now two quintessential ways in which information and knowledge can be shared: testimony and argumentation. Prima facie it may seem that, in cases of testimony, epistemic vigilance is directed mostly at the *source* of information, whereas in argumentation there is primarily engagement with the *content* communicated. This is indeed a plausible view, and one that aligns with much of what has been said about testimony and argumentation in the social epistemology literature.

Testimony occurs when someone tells something to another person, thus giving the second person the option of believing what she says. Whether testimony in general can be considered a legitimate source of knowledge is in fact an old philosophical question, which has attracted the interest of classical Indian philosophers (Chakrabarti and Matilal 1994) as well as early modern philosophers such as Locke (Boespflug 2019). Testimony has been extensively discussed in the epistemological literature in recent decades (Lackey and Sosa 2006) (Goldman and Blanchard 2015). In this literature, there are different accounts of the epistemology of testimony, but most of them agree that trust is a key component in testimony (Dormandy 2020). Indeed, a successful instance of testimony will typically be the result of the receiver judging the testifier to be sufficiently trustworthy so as to take their word at face value;

epistemic vigilance is directed primarily at the source of information (though the receiver may also perform some sort of coherence check on the content of the message).

But, as discussed in the literature on the norms of assertion, when witnessing someone engaging in an act of assertion, we typically take the asserter to have good (epistemic) reasons to make the assertion either because she knows its content to be true or because she has good reasons to think she is justified in that belief (Lackey J. 2007). Furthermore, in some cases, rather than simply assuming that the testifier must have good enough reasons to assert that P, we may also want to inspect those grounds ourselves, and thus we may ask for further support for the assertion. Indeed, as claimed by Brandom (1994) and taken up by a number of different authors, when someone makes an assertion that P he also thereby undertakes the obligation to justify that P upon request by others. That is, if challenged, the asserter must provide reasons for her assertion in what Brandom calls the "game of giving and asking for reasons." But whenever such a challenge is posed, that is, when further reasons are requested, we leave the domain of pure testimony and enter the domain of *argumentation*.

Crucially, when requesting reasons supporting an assertion that P, the receiver is no longer only directing epistemic vigilance towards the *source* as in the case of pure testimony. Instead, she is also engaging with the *content* of the claim. So it seems that, in this respect, testimony and argumentation can be seen as rather different mechanisms for epistemic vigilance: one is primarily source-oriented, while the other is primarily content-oriented. The question however is whether, in argumentation, epistemic vigilance towards the content will occur even in the absence of sufficient trust in the source. This is the main question to be addressed in this paper.

3. A Millian conception of argumentation

It is commonly thought that to engage in argumentation can be a way to acquire more accurate beliefs. By critically examining reasons for and against a given position, we would be able to weed out weaker, poorly justified beliefs (more likely to be false) and end up with stronger, suitably justified beliefs (more likely to be true). This

is indeed a very widespread view, perhaps even *the* most widely held view on argumentation among philosophers. On this view, argumentation is typically considered to be *truth-conducive*, at least in cases where there is an objectively correct answer to a problem. Goldman captures the gist of it in the following terms:

> Norms of good argumentation are substantially dedicated to the promotion of truthful speech and the exposure of falsehood, whether intentional or unintentional. [...] Norms of good argumentation are part of a practice to encourage the exchange of truths through sincere, non-negligent, and mutually corrective speech. (Goldman 1994, p. 30).

In this vein, a number of thinkers, most notably John Stuart Mill (and more recently [Betz 2013] and [Mercier and Sperber 2017]), maintain that interpersonal argumentative situations involving people who truly disagree with each other work best to realize the epistemic potential of argumentation to improve our beliefs.[2] Mill famously defended this position in *On Liberty* (1859) saying that when our ideas are challenged by engagement with those who disagree with us, we are forced to consider our own beliefs more thoroughly and critically.

> [Man] is capable of rectifying his mistakes, by discussion and experience. Not by experience alone. There must be discussion, to show how experience is to be interpreted. Wrong opinions and practices gradually yield to fact and argument; but facts and arguments, to produce any effect on the mind, must be brought before it. (Mill 1999, p. 41).

He describes this process as a *free exchange of ideas*. The expected result is that the remaining beliefs, those that have survived critical challenges, will be better grounded than those held before such encounters. As Mill puts it, "both teachers and learners go to sleep at

---

[2] I have defended this view myself (Dutilh Novaes 2021), but with the important caveat that the beneficial epistemic effect of interpersonal argumentation will come about only against the background of specific circumstances that ensure open-mindedness and good faith exchange of ideas. Moreover, the topic under discussion matters. For intellective problems, that is those that have a unique answer within a given theoretical framework (e.g. a mathematical or logical problem), group discussion has a clear beneficial, truth-conducive effect, but that is much less the case for so-called judgmental problems (Laughlin 2011).

their post, as soon as there is no enemy in the field"[3] (Mill 1999, p. 83).

Dissenters thus force us to stay epistemically alert instead of becoming too comfortable with existing, entrenched beliefs. But for this process to be successful, dissenters must be permitted to voice their opinions and criticism freely, and indeed Mill's forceful defense of free speech is one of his most celebrated positions. He emphasizes the role played by the free exchange of ideas in facilitating the growth of knowledge in a society: the more dissenting views and arguments in favor or against each of them are exchanged, the more likely it is that the "good" ones will prevail (Halliday and McCabe 2019).

But it is not sufficient that dissenters be given the opportunity to voice their opinions freely. It is also of crucial importance that receivers of these opinions and arguments be willing to engage with dissenters in good faith and with an open mind.[4] Mill pays much attention to the structural conditions for the free exchange of ideas but seems to fail to consider the well-documented tendency to avoid engaging with dissenting views altogether and to explain away contrary evidence so as to preserve prior beliefs.

Indeed, there is quite some evidence suggesting that arguments are in fact not a very efficient means to change minds in most real-life situations (Kolbert 2017) (Gordon-Smith, 2019). People typically do not like to change their minds about firmly entrenched beliefs they hold, so when they are confronted with arguments or evidence that contradict these beliefs, they tend to either look away or to discredit the source of the argument as unreliable—a phenomenon known as "confirmation bias" (Nickerson 1998). In particular, arguments that threaten our core beliefs and our sense of belonging to a group (e.g. political beliefs) typically trigger various forms of motivated reasoning whereby one outright rejects those arguments without engaging substantially with their content (Taber and Lodge 2006) (Kahan 2017). Engaging in argumentation is often a means to express and cement social identities rather than to come closer to

---

[3] Notice here an instance of the familiar "argument-as-war" conception of argumentation (Cohen 1995).

[4] There is also the question of whether dissenting voices will be paid attention to at all if they belong to marginalized groups. I return to this question briefly below.

the truth (Talisse 2019).

Relatedly, when choosing among a vast supply of options, people tend to gravitate towards content and sources that confirm their existing opinions in so-called "echo chambers" and "epistemic bubbles" (Nguyen 2020). These conversations with like-minded people may then reinforce their prior beliefs and even drive them to more extreme versions of those beliefs. In other words, the mere availability of dissenting voices is clearly not sufficient to ensure that knowers remain epistemically alert and consider all sides of a question, as suggested by Mill; there is always the possibility of simply not engaging constructively with these voices and the substance of their arguments in the spirit of a "free exchange of ideas" (or simply not engaging at all!). Moreover, the familiar phenomenon of polarization often arises: instead of bringing parties closer together, argumentation and deliberation often have the opposite effect of drawing them further apart (Sunstein 2017).

These observations suggest that we are not Millians when it comes to argumentation and dissent. The epistemic alertness that Mill believed would be the natural result of being exposed to dissenting opinions and arguments often fails to come about. The Millian account is thus descriptively inaccurate, at least with respect to many contexts (it might apply to some extent to some specialized contexts such as in science). One might retort that the Millian account is still *normatively* correct, but given the fact that it is arguably highly idealized, it is not suitable to offer prescriptive recommendations for fallible human agents. Instead, we need a more realistic approach to the (social) epistemology of argumentation, which will be sketched in the next sections.

4. Social exchange theory

In order to formulate a realistic account of argumentation as epistemic exchange, I draw inspiration from social exchange theory (SET). This is a framework developed by sociologists and social psychologists that seeks to explain human social behavior by means of processes of exchange between parties involving costs and rewards against the background of social networks and power structures (Cook 2013). SET is an influential and empirically robust

framework that has been used to investigate a wide range of social phenomena. The framework was originally developed in the late 1950s and early 1960s under the influence of work in economics (rational choice theory) and psychology (behaviorism), but it was also inspired by earlier anthropological work by Malinowski, Mauss, and Lévi-Strauss (Cook 2013). A key idea is that while humans have to protect their own interests to thrive, they are also highly dependent on others and on social structures more generally, which gives rise to an interplay between self-interest and interdependence. SET relies on concepts such as subjective cost-benefit analysis and comparison between alternatives while also emphasizing the importance of trust, reciprocity, and fairness (Stafford 2008).

In the words of sociologist Peter Blau, one of the pioneers in the field, "social exchange ... refers to voluntary actions of individuals that are motivated by the returns they are expected to bring and typically do in fact bring from others" (Blau 1964, p. 91). Social exchange is contrasted with formalized economic exchange where the exact terms of the exchange are made explicit beforehand. By contrast, social exchange "involves the principle that one person does another a favor, and while there is a general expectation of some future return, its exact nature is definitely not stipulated in advance." (Blau 1986, p. 93). Successful social exchanges can be mutually beneficial when both parties are better off by exchanging with each other rather than abstaining from it.

The concept of *resource* is a key component of SET as resources are what is exchanged in these interactions (Törnblom and Kazemi 2012). The concept is understood at a fairly high level of abstraction, which allows for different kinds of interactions to be conceptualized within the same model. For example, economic transactions, romantic relationships, workspace dynamics etc. can all be conceptualized as instances of social exchanges, but involving different kinds of resources. There are different taxonomies of resources available in the literature, but an influential taxonomy classifies resources as belonging to one of six categories: goods, status, affection, services, information, and money (Foa and Foa 1980). Importantly, exchanges may involve resources of different

kinds, for example when a service is provided in exchange for money.

Another key notion is that of an *exchange network* or *opportunity structure* defined as a set of possible or potential relations that limit and constrain who may exchange and interact with whom (Cook 2013). Typically, an agent will have a range of potential exchange partners to obtain a given resource that she deems valuable; she will then opt among these different options for the one that seems "best" (or least bad, when none of the options is truly advantageous) relative to a number of parameters (not only maximizing individual payoff, but also maintaining valuable social ties etc.). In other words, decisions to engage with a particular exchange partner are contrastive decisions: I choose to engage with A *instead of* engaging with B or C, with whom I could also have exchanged given the relevant opportunity structure. "[I]individuals form and maintain a relationship as long as the benefits from that relationship exceed those available elsewhere" (Lawler, Thye and Yoon 2008, p. 521).

Moreover, social exchanges can be viewed as belonging to different categories. Four main types of social exchanges are described in the literature (Molm, Collett and Schaefer 2007) (Lawler, Thye, and Yoon 2008, see figure 1):

- *Negotiated exchange*: actors jointly negotiate the terms of an agreement that benefits both parties either equally or unequally. Both sides of the exchange are agreed upon at the same time.
- *Reciprocal exchange*: actors perform individual acts that benefit another, such as giving assistance or advice, without negotiation and without knowing whether or when or to what extent the other will reciprocate. It is rather like a process of "gift-exchange."
- *Generalized exchange*: the recipient of benefit does not return benefit directly to the giver, but to another actor in the social circle. The giver eventually receives some benefit in return, but from a different actor.
- *Productive exchange*: interactions aimed at a jointly produced collective good wherein people unilaterally provide benefits to the group and receive benefit from it.
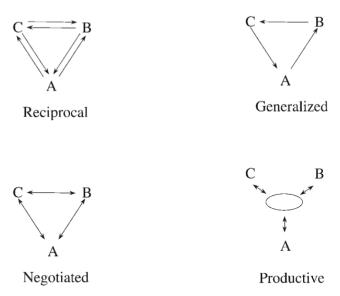
Figure 1: Four forms of social exchange (from (Lawler, Thye, and Yoon 2008)). The arrows indicate the flow of benefits from one agent to another.

Naturally, in all four forms of social exchange, but especially in reciprocal, generalized, and productive exchanges, a great deal of *trust* must be involved. Entering a social exchange is typically a trusting action, based on the expectation, but not the assurance, that the exchange partner will reciprocate at some point. Moreover, the resouces/benefits being offered must also be such that they are truly (rather than merely apparently) valuable for the receiver. The perspective adopted is diachronic, that is, of repeated exchanges, and the prediction is that, in the medium-long run, people will stop exchanging with given partners if these partners are repeatedly not "pulling their weight" as it were. "The failure of reciprocity results in infrequent exchange" (Cook 2013, p. 66). Thus, estimating the trustworthiness of different potential exchange partners is a fundamental component of choosing whom to exchange with; prior experiences with particular actors as well as reputational factors will be involved in these assessments.

This means that interpersonal trust is both an initial condition as well as a result of successful exchanges: actors will typically choose to exchange with partners that they consider more trustworthy than

the alternatives both in the sense that they offer valuable resources and in the sense that they will suitably reciprocate in the long run. Iterated successful exchanges will in turn increase trust in those partners, as well as positive affect towards them and towards the relationship as such (Lawler, Thye,and Yoon 2008).

Another well-established result in the SET literature is the correlation between *uncertainty* and *commitment*. Commitment refers to the extent to which an actor engages in repeated exchanges with the same partner over time. Uncertainty in turn refers to the probability of suffering from acts of opportunism imposed by one's exchange partners. "If an actor or subset of actors within a given opportunity structure prove themselves to be trustworthy, continued exchanges with that exchange partner provides a safe haven from other potentially opportunistic partners." (Cook 2013, p. 70).

Commitment thus understood seems like a laudable and reasonable principle, and indeed it is an essential factor for social cohesion. However, commitment also comes at a price: "such commitments, however, have the drawback of incurring sizable opportunity costs in the form of exchange opportunities foregone in favor of the relative safety provided by ongoing commitments" (Cook 2013, p. 70). In other words, if I continue to exchange with partner A because so far these exchanges have been largely successful from my perspective, this may preclude me from considering other options within the relevant opportunity structure. It may well be that exchanging with potential partner B instead would be even more advantageous, but by sticking to A exclusively, I forego this opportunity.

Unsurprisingly, a result observed both in lab experiments and in real-life scenarios is that actors have a greater tendency to form commitments in environments with high levels of uncertainty. These are environments where generalized trust towards potential partners is low on account of widespread occurrences of opportunism and a high frequency of actors acting in bad faith. In these scenarios, "actors were willing to forgo more profitable exchanges with untested partners in favor of continuing to transact with known partners who had demonstrated their trustworthiness in previous transactions" (Cook 2013, p. 71). Again, the rationale for commitment thus understood in situations of uncertainty is perfectly reasonable

in terms of the risk assessment necessary to choose among different exchange partners.

In this section, I've presented a very concise selection of concepts from social exchange theory, which will be relevant later on for the discussion of argumentation as epistemic exchange. The key points are:

- Engaging in an exchange is a voluntary action that consists of a choice to exchange with a given partner rather than with other potential partners based on the estimation that engaging with this particular partner will be more advantageous (or at least less disadvantageous) and/or less risky than engaging with any of the alternatives.
- What is exchanged in these interactions are different kinds of resources.
- Most of these exchanges will involve a certain amount of trust, as social exchange is generally speaking more like "gift-giving" and less like an economic transaction: there is expectation but no guarantee of reciprocity.
- In environments of (real or perceived) high uncertainty, actors will tend to continue to exchange with known partners, even if this will entail opportunity costs.

5. Argumentation as epistemic exchange

Now that some of the basics of SET have been laid out, let us further unpack the idea of argumentation as a form of epistemic exchange. Recall that Mill describes engaging with dissenters as a "free exchange of ideas." My claim here is that the reference to "exchange" is not purely metaphorical. Indeed, we can view social processes of circulation of knowledge, information, justification etc. as a particular kind of exchange, which we may describe as *epistemic exchange*. These exchanges will chiefly involve resources of a particular kind, which we may call *epistemic resources*: resources that are primarily involved in processes of production and circulation of knowledge.[5] Moreover, resources such as attention and time will

---

[5] We do not want to count, for example, solar light as an epistemic resource, even if it is also essential for the production and circulation of knowledge insofar it is

also be crucial for such processes. It may also happen that epistemic resources are exchanged for non-epistemic resources in a given interaction, for example a tutor who is paid to teach a pupil, but we may still view such exchanges as epistemic.

Thus understood, epistemic exchanges will be of various kinds. Argumentation is one specific type of epistemic exchange among others, including testimony. In argumentation, epistemic resources such as knowledge, evidence, justification, critical objections, defeaters can be exchanged by means of arguments. Thus, these resources will not only be of the kind that may expand one's set of beliefs, such as knowledge or justification, but also of the kind that may *decrease* one's set of beliefs, such as objections or defeaters.

Crucially, an argumentative exchange thus understood can be *mutually beneficial*; argumentation need not be a zero-sum game with winners and losers. True enough, the "argument as war" conception is widespread (as well as much criticized (Cohen 1995)), and there's no denying that, at least in some circumstances, people argue to win and to beat their "opponents" (Fisher and Keil 2016). But arguing to win is just one of the possible aims that arguers may have (Dutilh Novaes 2020). In contrast, when engaging in argumentation in the spirit of epistemic exchange, ideally both parties will come out as "winners" in the sense that they will have both benefited from the exchange. Simply put, they will both have learned something of value. Indeed, when the arguer offers the addressee some epistemic grounds supporting a given conclusion, she may be offering some valuable epistemic resources in the form of evidence, justification etc. that the addressee may benefit from. In turn, if the addressee suitably challenges the arguer's claims with appropriate objections, requests for further clarifications, or defeaters, this may lead to epistemic improvement for the arguer too in that she may be able to revise her initial beliefs (as indeed claimed by Mill).

Importantly, argumentation may also lead to the *joint production* of new epistemic resources. Recall that one of the four forms of

essential for human life in general. But solar light is not *primarily* involved in these processes. Admittedly, the concept of epistemic resource requires a much more detailed discussion than what I offer here. This concept is one of the topics of investigation that my research team is currently working on.

social exchange discussed in the previous section is *productive exchange*: interactions aimed at a jointly produced collective good. A paradigmatic example would be the social epistemology of mathematical proofs. As described by Lakatos (1976), in terms of a dialectic of proofs and refutations, mathematical proofs are typically a joint effort involving those proposing a proof and those inspecting the proof for possible mistakes and counterexamples, or simply for steps in the proof that are not sufficiently clear. These "critics" may be colleagues, or students, or, more formally, the referees who are asked to peer-review papers for journals (Andersen 2018) (Dutilh Novaes 2021). This critical engagement is an integral part of the process of formulating new proofs, so much so that the "voice" of the critics/referees is typically still subtly present in the end-result of the proof. And while the case for productive exchange through argumentation may be the clearest in mathematics, it is arguably also present in other (scientific) domains, especially given the role of the peer review system. The protocol of "adversarial collaboration" in the social sciences (Mellers, Hertwig, and Kahneman 2001) would be a particularly clear example of productive exchange through argumentation.

6. Can argumentation compensate for the absence of sufficient trust?

Let us now return to the concept of epistemic vigilance and to the point that argumentation may allow for the transmission of content even when there is diminished trust between the parties. After all, this would (presumably) be what differentiates argumentation from testimony. "The exchange of arguments improves communication by allowing messages to be transmitted even in the absence of sufficient trust" (Mercier 2016, p. 690). The thought is that, with successful testimony, the content of a message is accepted without reservation because the receiver trusts the sender; no further reasons are required. (Though the receiver may run a coherence check on her prior beliefs.) In the absence of trust, the receiver will require additional reasons to be convinced, for example, by engaging with the content of the message critically and requesting additional

support for the claim. This is precisely when the "game of giving and asking for reasons" that constitutes argumentation will kick in.

However, if argumentation is indeed suitably conceived as a form of exchange, and if social exchange theorists are correct in claiming that evaluations of trustworthiness play an important role in deciding whom to engage with in exchanges, then the prediction will be that argumentation will *not* necessarily facilitate the transmission of messages in the absence of sufficient trust. To enter an argumentative situation in the spirit of honest epistemic exchange (especially when there is disagreement), a person must think that the reasons offered by other parties are minimally reliable. For example, if she suspects that someone presenting arguments is in fact deliberately trying to mislead her, she may reasonably assume that the arguments being offered are epistemically defective, as in the case of a dishonest seller trying to sell a malfunctioning car.

The role of trust in argumentation and epistemic exchanges more generally becomes even more palpable if one considers the key idea from SET that exchanges occur against the background of *opportunity structures*, where actors typically have the choice among different possible exchange partners. Naturally, they will tend to choose those who they deem to be more trustworthy than others (though this is just one of the factors influencing such choices, albeit a crucial one). The same holds for epistemic exchanges: given limitations of time and attention, it makes sense to focus on those partners who the actor has reason to believe have valuable epistemic resources to offer[6] and are not acting in bad faith. For this to be the case, these partners must not only be sufficiently competent/knowledgeable on the topic in question, but they must also not be acting dishonestly or opportunistically.

Thus, in this respect, the conceptualization of argumentation as epistemic exchange being articulated here differs substantially from the Sperber/Mercier epistemic vigilance model in that it postulates

---

[6] Naturally, this may give rise to occurrences of epistemic injustice, as the estimation of how valuable a potential partner is for epistemic exchange may be influenced by epistemically irrelevant factors such as gender or race (more on epistemic injustice shortly).

that a certain amount of trust is typically (though perhaps not always) required for an argumentative exchange to be epistemically fruitful. This means that, contrary to what Mercier claims, argumentation will often *not* be a solution to the problem of communication in the absence of sufficient trust.

Another line of argument supporting the hypothesis that trust plays a significant role in argumentation comes from K. Hawley's work on trust and trustworthiness, in particular as pertaining to assertions (Hawley 2014; 2019 Chapter 3). She writes:

> Trusting someone to do something involves both trust in her competence and trust in her willingness to act; part of trustworthiness is the attempt to avoid commitments you are not competent to fulfill. The special case of trusting what someone says involves both trusting her to possess the truth (or knowledge), and trusting her to speak sincerely. People can betray our trust through either error or deceit, and part of trustworthiness is the attempt to avoid making assertions where you lack knowledge… (Hawley 2014, p. 17).

So far no surprises here, as the significance of trust and trustworthiness in the context of testimony/assertion has already been recognized. But do these observations on *assertion* carry over to *argumentation*? In fact, it is plausible to view an argument as *a special kind of assertion*, or as a structured collection of assertions. For example, in a categorical argument of the form 'A, therefore B,' three assertions seem to be made: A, B, *and* that there is a relevant relation of support between A and B (by contrast, in a hypothetical argument of the form 'suppose A, then B would be the case,' the only assertion made is that there is a relevant relation of support between A and B). A precursor of this view is Frege's observation that an inference is "the pronouncement of a judgement made in accordance with logical laws on the basis of previously passed judgments" (Frege 1984, p. 318) For Frege, inference is itself a judgment, which takes judgments (premises) to judgment (conclusion). (Frege's inferences would correspond to arguments, and Frege's judgments to assertions in the present framework.)

What is the epistemic status of arguments relative to the epistemic status of assertions for receivers? One might think that arguments should be epistemically more transparent than assertions, as the receiver can inspect the relation of support between premises

and conclusion herself. But just as with factual assertions of the kind "Miss Scarlet killed the butler," where the receiver is often not in a position to ascertain its truth herself and must thus rely on the testimony of an informant, the receiver of an argument is often *not* in a position to assess the cogency of the relation of support between premises and conclusion. This may be the case in virtue of domain-specific knowledge that is required to evaluate the argument, but also simply because, contrary to what many of us would prefer to believe, correctly evaluating the cogency of arguments is in fact a rather difficult task.[7] As widely documented by research into reasoning biases, reasoners are often led astray in their evaluations of arguments due to the influence of a number of prima facie irrelevant factors (Stanovich 2003). The very fact that, for millennia, argumentation theorists and logicians have spent considerable time and energy identifying and discussing *fallacies* is another indication of how easy it is to be "fooled" by arguments that appear correct but are not. If receivers of arguments have some level of meta-cognitive awareness of their own limitations and vulnerability to misleading arguments (which is, admittedly, an empirical question), then one should expect that trustworthiness will be a concern for the receiver of arguments just as it is for the receiver of assertions: she wants to maximize the likelihood that the epistemic resources being offered are indeed of good quality.

I have presented two arguments supporting the claim that trust and the trustworthiness of the source are significant factors for receivers of arguments when evaluating their correctness. One is based on the idea that argumentation is a form of epistemic exchange and that all exchanges involve evaluations of trustworthiness; and the second one is based on an epistemic similarity between assertions and arguments, namely that receivers are not always in a good position to evaluate their truth/correctness.[8] From an exchange perspective, there are two senses in which trust is relevant: for the giver, the question is whether the receiver will

---

[7] But note that, in a recent book, Mercier (2020) argues that we are in fact well equipped to ascertain what and whom to believe. By contrast, other researchers argued vehemently that we are easily "duped" (Levine 2019).
[8] Yap (2013) has pointed out other ways in which credibility plays a role in argumentation similar to its role in testimony.

reciprocate in due course; for the receiver, the question is whether the resources being offered by the giver are 'of good quality' and will indeed be valuable and beneficial for the receiver.[9] The present discussion has focused on the receiver's side,[10] and the question for the receiver is primarily whether the argument being offered is of good quality. Hawley's account of trusting someone who is asserting suggests two ways in which an an arguer may betray the trust of an addressee by offering a low quality argument. This may happen if the arguer is simply not sufficiently knowledgeable on the topic at hand and thus makes a mistake; or if the arguer is deliberately trying to deceive the receiver, for example by presenting an argument to "trick" her into doing something that is not in her best interest.

Moreover, not only can argumentation not "fix" communication in the absence of sufficient trust: argumentation may even *erode* trust, i.e. lead to a decrease in trust between the communicating parties. Two relevant pieces of evidence can be offered for this claim. Firstly, in the literature on peer disagreement, it has been noted that, if a presumed peer holds a view that appears to be simply preposterous from the perspective of a given agent, a rational response by this agent is to *downgrade* the other person from her status as a peer (Worsnip 2014).[11] Secondly, work on Bayesian simulations of group deliberation shows that, when agents initially trust each other but hold different views, if they engage in argumentation/deliberation they in fact come to trust each other *less* than before (Olsson 2013). The conclusion is thus that erosion of trust is one of the risks/potential losses that may be incurred as a result of engaging in argumentation, which thus must enter the

---

[9] Of course, trust is a multifaceted phenomenon, which manifests itself in many ways. Here I have focused on interpersonal trust, and in particular on these two manifestations: expectations of reciprocity, and expectations regarding the quality and value of the resources being offered (e.g. whether a seller is not offering a malfunctioning car under the guise of a well-functioning one).

[10] For the giver/arguer, one important question is whether a potential receiver is likely to truly listen and pay attention to the argument. If not, then it would be a waste of time to engage in argumentation with an unresponsive addressee.

[11] To be fair, this is a much disputed issue in the peer disagreement literature, and many authors claim that the "conciliatory" response, whereby one moves closer to the dissenting view of the peer, is the more rational response.

"cost-benefit analysis" on whether to argue or not (Paglieri and Castelfranchi 2010).

To be clear, the claim is not that trust is a *necessary* condition for argumentation to come about. In many circumstances (e.g. political debates), one will have good reasons to engage in argumentation with someone one does not trust (in some relevant sense of "trust"). The claim is rather that, when argumentation is meant to be a form of epistemic exchange from which one hopes to learn something of value, trust and trustworthiness do play a role, and an arguer will typically choose to engage with a partner that she considers to be more trustworthy than the alternatives (though perhaps not always, as trust is one among a number of factors to be taken into account). This is the case both descriptively and prescriptively speaking, given cognitive limitations of the arguer and the likelihood of opportunistic behavior in exchange partners.

These observations lead to the formulation of a three-tiered model of epistemic exchange and argumentative engagement, which draws substantially from social exchange theory.

1. **Attention/exposure**. The first stage of argumentative engagement consists in establishing whether two people are potential exchange partners at all, given the relevant opportunity structure for epistemic engagement. In simpler terms: who is in an agent's network of potential contacts? Who is in a position to attract someone's attention? It may be that potential lines of communication are cut, say in the case of structural censorship of some messages/senders or echo chambers (Nguyen 2020). At the other end of the spectrum, it may be that so many signals are being broadcast that the many different sources are all competing for the receiver's attention, in a so-called "attention economy" (Franck 2019).[12]

2. **Choosing whom to engage with**. The next stage involves making choices against the background of possibilities for exchange as determined by the relevant opportunity structure. Typically, there will be a number of options, and

---

[12] See (Dutilh Novaes and de Ridder forthcoming.) for a discussion of scarcity vs. overabundance of information in an epistemic environment.

given limitations of time and attention, choices will have to be made also in the context of epistemic exchanges (even if it is reasonable to expect that one will engage with a wider range of partners for epistemic exchanges than when one buys a car, for example). Among those who have caught my initial attention, whom do I further engage with? It is at this point that considerations of trustworthiness come into play (among other factors). In particular, choosing to trust someone will sometimes entail choosing *not* to trust someone else, especially when their respective messages conflict.

3. **Engagement with content**. It is only at a third stage that engagement with the *content* of the argument properly speaking should occur. This is when the actual epistemic exchange in fact takes place. At this point the receiver will reflectively (and perhaps critically) engage with the argument seeking to understand its substance and evaluate its cogency. In case of a positive evaluation, this may lead to a change in view for the receiver. It may also lead to a mutually beneficial exchange where both arguer and addressee improve their respective epistemic stances and in some cases even go on to create new epistemic resources together (productive exchange).

One of the advantages of this model is that it offers resources to account for various forms of epistemic injustice. Regarding the first stage, it highlights a familiar problem for marginalized groups who are pushed towards the fringes of exchange networks. Effectively, they are not part of the "conversation," as it were. What good is it to be voicing arguments against oppression, if no one who might be able to do something about it is actually paying attention? Regarding the second stage, comparative judgments of trustworthiness may give rise to forms of epistemic injustice if they are motivated by epistemically irrelevant factors such as gender or race (Medina 2011). Finally, in the third stage forms of argumentative injustice may arise if the social identity of the arguer influences judgments of the cogency of the argument by receivers (Bondy 2010); the bar for persuasion may be set higher for arguers

that are viewed as prima facie less competent in virtue of gender or race etc.

Note that this three-tiered model contrasts with models where epistemic evaluating of the source and the content of a message happen in parallel, and/or with significant feedback loop between the two (which I take to be the gist of the Mercier-Sperber epistemic vigilance model). In the next section, I briefly discuss two case studies that seem to be instances of the three-tiered model rather than of alternative models.

7. Some examples

So let us now discuss some concrete examples of argumentation in real-life situations so as to gauge the soundness of the three-tiered model described above and the hypothesis concerning the role of trust in argumentation. The two examples to be discussed are vaccination hesitancy and authoritarian politicians who artificially create a perception of uncertainty towards traditional sources of information such as journalism and science.

Vaccination hesitancy was considered to be one of the top ten threats to global health by the World Health Organization in 2019.[13] As is well known, it is a tendency that has been intensified in the last decades, in particular as a result of an infamous and now retracted paper claiming that there is a connection between vaccination and the onset of autism. The puzzling thing is, even now that this claim has been thoroughly debunked, vaccination hesitancy remains strong, which seems to suggest that purely scientific arguments are not sufficient to dispel the persistent perception of a high likelihood of negative outcomes as a result of vaccinating children. Indeed, it has been shown that these so-called "vaccination myths" are not easily debunked with scientific arguments (Horne, Powell, Hummel, and Holyoak 2015). Even more disheartening: other fact-based communication strategies and interventions do not seem to fare much better either (Nyhan, Reifler, Richey, and Freed 2014).

---

[13] https://www.who.int/news-room/feature-stories/ten-threats-to-global-health-in-2019

What could explain the obstacles to changing hesitant minds on the safety of vaccination with scientific arguments? A plausible explanation is the high prevalence of mistrust towards science in general and the pharmaceutical industry in particular. And the truth is, this mistrust is oftentimes not entirely unwarranted. A 2019 editorial in *The Lancet Infectious Diseases* entitled "Trust Issues" addressed precisely this problem:

> In the USA, the country is plagued by prescription opioid misuse fuelled by aggressive pharmaceutical marketing, the people of Flint, MI, have been without safe drinking water for 3 years, and the most basic drugs are often unaffordable because of profit-driven health care. Little wonder that some individuals question the authorities' desire to prioritise their wellbeing. It is impossible to build trust while at the same time abusing it.[14]

While mistrust in "Big Pharma" may be more pronounced in the USA than in other countries, this is a general problem that affects vaccination compliance elsewhere too. In fact, a 2018 systematic review of the literature on vaccination and trust (Larson, et al., 2018) found that, while more research is needed to understand trust as it relates to vaccination acceptance, it seems that people are more likely to get vaccinated if they trust the government, their health care system, their health care provider, or other members of society in general. One factor that seems to have a decisive effect in people's decisions to vaccinate or not is their relationship with their primary health care provider. As aptly described in a journalistic piece on trust and vaccination:

> [T[he one point in the system that everyone seems to agree has the most potential for building trust is the doctor-patient relationship. "Most of our interviewees trust their personal doctors a lot," said Amelia Jamison, a faculty research assistant the University of Maryland Center of Health Equity who studied trust related to the flu vaccine. "People are more likely to accept vaccines when their health care providers strongly recommend them," Jamison said. It also helps, she added, when doctors act as role models, and

---

[14] https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(19)30128-8/fulltext

explain to patients why they choose to vaccinate themselves and their own families.[15]

Thus, vaccination hesitancy and resistance to scientific arguments supporting vaccination seem to emerge as a fine example of the three-tiered model of argumentative engagement sketched above. Once trust in certain kinds of sources—traditional scientific sources in this case—is severely undermined, the content of the arguments themselves becomes irrelevant from the perspective of the mistrustful receiver, and there is no further engagement. By contrast, where trust is in place or can be built again, in particular in the doctor-patient relationship, vaccination hesitancy can be overcome.

The second example is the familiar tactic often adopted by authoritarian politicians to discredit sources of information that they cannot control, such as traditional journalism and media, or scientists and other experts (Stanley 2018). Trump's infamous appropriation of the term "fake news" to undermine trust in mainstream media is a recent example of this phenomenon, which has a clear precursor in the term "Lügenpresse" ('lying press') as used by Hitler and his supporters in the 1930s (Snyder 2018).

Here the observation from SET that environments of high (perceived or real) uncertainty give rise to more commitment is illuminating. When told that other sources of information cannot be trusted, which creates a perception of uncertainty, it is not surprising that a consumer of news who in fact "buys" into this narrative is going to focus his attention only on the news sources that he is told are trustworthy; he will *commit* to these sources, in the sense understood by social exchange theorists. This is the proverbial "Fox News viewer" who only consumes news from certain sources, thus becoming encapsulated in a right-wing media environment (Benkler, Faris, and Roberts 2018), among other reasons because he is made to believe that all other sources have hidden agendas and engage in systematic lying. Even if the percentage of people who indeed fit this profile is ultimately not as large as one might think (as has been argued),[16] it is still the case that a politician like Trump seeks to

---

[15]https://undark.org/2019/04/24/anti-vaxxers-vaccines-trust-big-pharma/
[16]https://fivethirtyeight.com/features/media-bubbles-arent-the-biggest-reason-were-partisans/

exploit the connection between commitment and uncertainty by artificially creating a perception of uncertainty. How successful this strategy is going to be depends on a number of factors, some of them beyond the politician's control (thankfully!).

Here again we note that commitment in repeated exchanges, while surely to some extent rational in the context of (real or perceived) uncertainty, entails *opportunity costs*. Those who consume media from a narrow range of sources are missing out on the opportunity to come into contact with a broader spectrum of views, arguments and information, from which she might have otherwise benefited—that is, if her goal is indeed the epistemic goal of acquiring more accurate, better grounded beliefs (of course, her goal may well *not* be primarily epistemic in the first place).

Notice that, when someone is selecting which media sources she will pay attention to and engage with, what is taking place are the first two stages described in the three-tiered model of argumentative engagement described above. How much effort will she put into expanding the possible range of sources, i.e. her opportunity structure for epistemic exchange? (Stage 1) And given the alternatives presented to her, on what basis will she choose the sources to engage with among these options, in particular given limitations of time and attention? (Stage 2) Moreover, while it is of course preferable if an agent can make autonomous choices at these stages, there is much that those in position of power can do to interfere with these processes (Dutilh Novaes and de Ridder forthcoming). This may happen either by reducing the range of options for epistemic exchange by "brute force" (through censorship for example), or by distorting perceptions of trustworthiness by, in the words of propagandist S. Bannon, "flooding the zone with shit."[17]

8. Conclusions

In this paper, I have argued that, contrary to what may be thought and is sometimes explicitly claimed, argumentation is not likely to be a solution to the problem of transmission of messages in the

---

[17] https://www.vox.com/policy-and-politics/2020/1/16/20991816/impeachment-trial-trump-bannon-misinformation

absence of sufficient trust. This is because an agent will prefer to engage in epistemic exchange by means of argumentation with partners whom she trusts, at least above a certain threshold. I've argued for this claim on the basis of a conceptualization of argumentation as epistemic exchange that borrows extensively from social exchange theory. I have then discussed two real-life examples that seem to illustrate basic principles from social exchange theory in argumentative situations: the debates around vaccination and vaccination hesitancy, and the undermining of the credibility of traditional sources of information by some authoritarian politicians.

But if it cannot compensate for the absence of sufficient trust, does this mean that argumentation becomes useless? After all, where there is sufficient trust, perhaps testimony alone suffices for the transmission of content, in which case argumentation becomes superfluous. But I think we can resist this pessimistic conclusion regarding the purported "uselessness" of argumentation. Engaging in argumentation may bring about a number of epistemic benefits, in particular related to *understanding*, which is an epistemic state related to but distinct from knowledge. As an approximation, understanding can be described as "a grasp of a comprehensive body of information that is grounded in fact, is duly responsive to evidence, and enables non-trivial inference, argument, and perhaps action regarding that subject the information pertains to" (Elgin 2007, p. 39). The fact that both understanding and argumentation are crucial aspects of scientific practice in particular seems to reinforce the hypothesis that one of the main epistemic functions of argumentation pertains to enhancing understanding by means of the exchange of epistemic resoures. The exact details of this putative connection between understanding and argumentation remain at this point a topic for future research. For now the suggestion is that the social epistemic import of argumentation may well pertain not to compensating for the absence of sufficient trust, but rather to some other epistemic phenomenon, for example understanding.

Be that as it may, the main take-home message from the present investigation is perhaps that argumentation theorists should not only focus on the argumentative exchanges that in fact take place and on the content of these exchanges. It is also crucial to investigate the stages prior to actual exchanges, in particular the choices that an

agent makes on whether to engage or not in argumentation, and on whether to engage in argumentation with interlocutor A rather than interlocutor B. For these choices, trust is an essential factor.

Acknowledgments

References

Andersen, L. E. 2018. Acceptable gaps in mathematical proofs. *Synthese, Online First*: 1-15.

Bell, D. E., Raiffa, H., and A. Tversky. 1988. Descriptive, normative, and prescriptive interactions in decision making. In *Decision making: Descriptive, normative, and prescriptive interactions*, eds. D.E. Bell, H. Raiffa, and A. Tversky, 9-30. Cambridge: Cambridge University Press.

Benkler, Y., Faris, R., and H. Roberts, H. 2018. *Network Propaganda.* New York, NY: Oxford University Press.

Betz, G. 2013. *Debate Dynamics: How Controversy Improves Our Beliefs.* Dordrecht: Springer.

Blau, P. M. 1964. *Exchange and Power in Social Life.* New York, NY: Wiley.

Blau, P. M. 1986. *Exchange and power in social life (2nd printing).* New Brunswick, NJ: Transaction Books.

Boespflug, M. 2019. Locke on testimony. *British Journal for the History of Philosophy* 27: 1135-1150.

Bondy, P. 2010. Argumentative Injustice. *Informal Logic* 30: 263-278.

Brandom, R. 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment.* Cambridge, MA: Harvard University Press.

Chakrabarti, A., and B. Matilal. 1994. *Knowing From Words: Western and Indian Philosophical Analysis of Understanding and Testimony.* Dordrecht: Kluwer.

Cohen, D. H. 1995. Argument is War...and War is Hell: Philosophy, Education, and Metaphors for Argumentation. *Informal Logic* 17: 177-188.

Cook, K. S. 2013. Social exchange theory. In *Handbook of social psychology,* eds. J. DeLamater, and A. Ward, 6-88. Springer.

Dormandy, K. 2020. Introduction: An Overview of Trust and Some Key Epistemological Applications. In *Trust in Epistemology* ed. K. Dormandy, 1-40. London: Routledge.

Dutilh Novaes, C. 2020. Reassessing different conceptions of argumentation. In *Polarisation, Arrogance, and Dogmatism: Philosophical Perspectives,* eds. A. Tanesini, and M. Lynch, London: Routledge.

Dutilh Novaes, C. 2021. *The Dialogical Roots of Deduction.* Cambridge: Cambridge University Press.

Dutilh Novaes, C., and J. de Ridder. (forthcoming.). Is fake news old news? In *The Epistemology of Fake News*, eds. S. Bernecker, T. Grundmann, and A. Flowerree. Oxford: Oxford University Press.

Easwaran, K. 2009. Probabilistic proofs and transferability. *Philosophia Mathematica* 17: 341-362.

Elgin, C. 2007. Understanding and the facts. *Philosophical Studies 132*: 33-42.

Fisher, M., and F. Keil. 2016. The trajectory of argumentation and its multifaceted functions. In *The Psychology of Argument: Cognitive Approaches to Argumentation and Persuasion* ed. F. Paglieri, 347-362. London: College Publications.

Foa, E. B., and U.G. Foa. 1980. Resource theory. In *Social exchange* eds. K. J. Gergen, 77-94. New York, NY: Springer.

Franck, G. 2019. The economy of attention. *Journal of Sociology* 55: 8-19.

Frege, G. 1984. On the Foundations of Geometry: Second Series. In *Collected Papers on Mathematics, Logic, and Philosophy*, 293–340. Oxford: Blackwell.

Goldman, A. I. 1994. Argumentation and social epistemology. *Journal of Philosophy* 91: 27-49.

Goldman, A. I. 1997. Argumentation and interpersonal justification. *Argumentation* 11: 155-164.

Goldman, A. I. 2004. An Epistemological Approach to Argumentation. *Informal Logic* 23: 49-61.

Goldman, A., and T. Blanchard. 2015. Social Epistemology. In *The Stanford Encyclopedia of Philosophy*, ed. E. Zalta: http://plato.stanford.edu/entries/epistemology-social/

Gordon-Smith, E. 2019. *Stop Being Reasonable: How we Really Change Minds.* New York, NY: Public Affairs.

Greenberg, M. S. 1980. A Theory of Indebtedness. In *Social exchange* ed. K. Gergen, 3-26. New York, NY: Plenum Press.

Halliday, D., and H. McCabe. 2019. John Stuart Mill on Free Speech. In *The Routledge Handbook of Applied Epistemology* eds. D. Coady, and J. Chase, 72-87. London: Routledge.

Harris, P. L. 2012. *Trusting what you're told.* Cambridge, MA: Harvard University Press.

Hawley, K. 2014. Trust, distrust and commitment. *Nous* 48: 1-20.

Hawley, K. 2019. *How To Be Trustworthy.* Oxford: Oxford University Press.

Horne, Z., Powell, D., Hummel, J. E., and K. J. Holyoak. 2015. Countering antivaccination attitudes. *Proceedings of the National Academy of Sciences of the United States of America* 112: 10321–10324.

Kahan, D. 2017. *Misconceptions, Misinformation, and the Logic of Identity-Protective Cognition.* Yale Law School, New Haven.

Kolbert, E. 2017. (February 27). Why Facts Don't Change Our Minds. *The New Yorker* .

Lackey, J. 2007. Norms of assertion. *Nous* 41: 594-626.

Lackey, J., and E. Sosa. 2006. *The Epistemology of Testimony.* Oxford: Oxford University Press.

Lakatos, I. 1976. *Proofs and Refutations.* Cambridge: Cambridge University Press.

Larson, H. J., Clarke, R. M., Jarrett, C., Eckersberger, E., Levine, Z., Schulz, W. S., et al. 2018. Measuring trust in vaccination: A systematic review. *Human vaccines & immunotherapeutics* 14: 1599–1609.

Laughlin, P. R. 2011. *Group problem solving.* Princeton, NJ: Princeton University Press.

Lawler, E. J. 2001. An Affect Theory of Social Exchange. *American Journal of Sociology* 107: 321-352.

Lawler, E. J., Thye, S. R., and J. Yoon. 2008. Social exchange and micro social order. *American Sociological Review* 73: 519–542.

Levine, T. 2019. *Duped: Truth-Default Theory and the Social Science of Lying and Deception.* Tuscaloosa: University of Alabama Press.

Medina, J. 2011. The Relevance of Credibility Excess in a Proportional View of Epistemic Injustice: Differential Epistemic Authority and the Social Imaginary. *Social Epistemology* 21: 15-35.

Mellers, B., Hertwig, R., and D. Kahneman. 2001. Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science* 12: 269-275.

Mercier, H. 2016. The argumentative theory: Predictions and empirical evidence. *Trends in Cognitive Science* 20: 689-700.

Mercier, H. 2020. *Not Born Yesterday.* Princeton, NJ: Princeton University Press.

Mercier, H., and D. Sperber. 2017. *The Enigma of Reason.* Cambridge, MA: Harvard University Press.

Messick, D., and K. Cook. 1983. *Equity theory: psychological and sociological perspectives.* Westport: Praeger.

Mill, J. S. 1999. *On Liberty.* Peterborough: Broadview Press.

Molm, L., Collett, J., and D. Schaefer. 2007. Building solidarity through generalized exchange: A theory of reciprocity. *The American Journal of Sociology* 113: 205–242.

Nguyen, C. T. 2020. Echo chambers and epistemic bubbles. *Episteme* 17: 141-161.

Nickerson, R. S. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*: 175- 220.

Nyhan, B., Reifler, J., Richey, S., and G. Freed. 2014. Effective messages in vaccine promotion: a randomized trial. *Pediatrics* 133: e835-e842.

Olsson, E. J. 2013. A Bayesian Simulation Model of Group Deliberation and Polarization. In *Bayesian Argumentation: The practical side of probability* ed. F. Zenker, 113-333. Dordrecht: Springer.

Paglieri, F., and C. Castelfranchi. 2010. Why argue? Towards a cost–benefit analysis of argumentation. *Argument & Computation* 1: 71-91.

Snyder, T. 2018. *The Road to Unfreedom.* London: Vintage.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., and G Origgi. 2010. Epistemic Vigilance. *Mind and Language* 25: 359-393.

Stafford, L. 2008. Social Exchange Theories. In *Engaging theories in interpersonal communication: Multiple perspectives,* eds. L. A. Baxter, and D. O. Braithwaite, 377–89. Thousand Oaks.

Stanley, J. 2018. *How Fascism Works.* New Haven, CO: Yale University Press.

Stanovich, K. E. 2003. The Fundamental Computational Biases of Human Cognition: Heuristics That (Sometimes) Impair Decision Making and Problem Solving. In *The Psychology of Problem Solving* eds. J. Davidson, and R. J. Sternberg, 291–342. Cambridge: Cambridge University Press.

Sunstein, C. (2017). *#Republic: Divided Democracy in the Age of Social Media.* Princeton, NJ: Princeton University Press.

Törnblom, K., and A. Kazemi. 2012. Some conceptual and theoretical issues in Resource Theory of Social Exchange. In *Handbook of Social Resource Theory: Theoretical Extensions, Empirical Insights, and Social Applications* eds. K. Törnblom, Aand. Kazemi, 33-64. New York, NY: Springer.

Taber, C. S., and M. Lodge. 2006. Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science* 50: 755–69.

Talisse, R. 2019. *Overdoing Democracy.* Oxford: Oxford University Press.

Van Lange, P. A., and D. Balliet. 2015. Interdependence theory. In *APA Handbook of Personality and Social Psychology, Vol. 3. Interpersonal Relations,* eds. M. Mikulincer, P. R. Shaver, J. A. Simpson, and J. F. Dovidio, 65-92. Washington, DC: American Psychological Association.

Worsnip, A. 2014. Disagreement about Disagreement? What Disagreement about Disagreement? *Philosophers' Imprint* 14: 1-20.

Yap, A. 2013. Ad Hominem Fallacies, Bias, and Testimony. *Argumentation* 27 (2):97-109.