Do we Trust Blindly on the Web?

Genot, Emmanuel; Olsson, Erik J

Published in:
Iride: Journal of Philosophy and Public Debate

DOI:
10.1414/86523

2017

Document Version:
Early version, also known as pre-print

Do we Trust Blindly on the Web? Origgi's Puzzle of Online Interaction and its Empirical Basis

Emmanuel J. Genot and Erik J. Olsson

Abstract: Gloria Origgi argues that our use of online search engines and online social network gives rise to a puzzle of trust: trust online is "blind", "massive" and "naïve" while this is not so for trust offline. Origgi's explanation is that offline trust is "relational" in the sense of being based on rational choice in a cooperative environment whereas online trust is "epistemic" and not based on rational choice in such an environment. We focus on Origgi's intriguing hypothesis about online trust and, more specifically, about the empirical basis for her characterization of such trust. For this purpose, we review relevant recent literature on online behavior. While we find much to agree with in Origgi's analysis, some of her testable claims are only partially supported by empirical evidence and there is, in some cases, evidence to the contrary. We conclude that the empirical case for online trust being generally epistemic is not compelling and, except in a few special cases, trust online is best described as being relational and strategic in character.

1. Introduction

Online social networks (OSN) and online search engines (OSE) are not only ubiquitous features of our daily life, but also a subject for scholarly research. Internet users' trust, or lack thereof, relative to both OSE and OSN, is an important topic in this research. Trust is involved, for instance, in how Internet users evaluate the credibility of OSE search results and the content they access and/or share in OSN, as well as in their decisions to share information in presence of privacy concerns or risks of cyber-bullying. A better understanding of the origins and nature of online trust can thus help answer questions relative to a wide variety of online behavior, from information-seeking to information-sharing.

Arguably, most approaches to online trust adopt implicitly or explicitly a cost-benefit analysis, which views an internet user as a decision-maker with expectations based on her beliefs and goals based on

her preferences. Such an agent takes one-time or sequential decisions based on her current expectations and goals. In the particular context of online behavior and interaction, these goals can be acquiring information, obtaining social gratification etc. The corresponding decisions are: take at face value or not the information accessed through an OSE or OSN, or share or not some content on an OSN. When decisions are sequential, the agent can update her beliefs based on the observed consequences of her past decisions, form new expectations and update her strategies accordingly. For instance, if sharing has been rewarded in the past, the user will continue to share.

Gloria Origgi (2013), noted for her influential philosophical work on reputation and trust,[1] has challenged the relevance of this cost-benefit approach. According to her, online trust is "the most naïve and wild form of blind trust that we have ever experienced in mature societies" (Origgi, 2013, p.25). This gives rise to a "puzzle of trust": how can it be that people, when they are online, suddenly start to act in this seemingly irrational manner? In order to solve this puzzle, Origgi appeals to a distinction between relational trust and epistemic trust. While the former is compatible with informed decisions whether or not to trust, the latter is not, representing as it does a "blind buy-in". According to Origgi, "[t]he web is a 'trust machine' that feeds itself with [. . .] social information". Subsequently, the very structure of the internet elicits epistemic trust, so that Internet users adopt a trustful attitude by default.

Even so, Origgi believes that it is possible as well as desirable for people to adopt a more cautious attitude toward online information. She argues, along with others, that evolution has endowed humans with an aptitude for epistemic vigilance (Sperber et al., 2010). This "vigilant attitude" can be activated in an online context by making people more sensitive to online cues of reliability or deception and generally by improving the digital literacy in society.

In section 2 we present Origgi's puzzle of trust, and in section 3 we identify a number of testable empirical claims in her account of the puzzle. These claims are, in section 4-6, subjected to empirical scrutiny based on the recent relevant literature on online interaction. We show that Origgi's claims

[1] For instance, Origgi (2004), Origgi (2005) and Origgi (2012).

about OSE are largely vindicated by empirical research but that the support for her claims about OSN is generally less convincing and in some cases non-existing, and that some of her claims about OSN are even false. Finally, we observe that Origgi's remarks about there being room for greater online vigilance are firmly in line with those made by web researchers on both OSE and OSN.

2. Origgi's puzzle of trust

The starting point of Origgi's 2013 article is a characterization of trust online as "disenchanted" (p. 23):[2] "social relations in late modern democracies are characterized by a form of disenchanted trust, that is, trust that comes out of a series of procedures of taming distrust, such as contracts, law enforcements, transparent procedures (concerning vote, attribution of rights, allocation of resources, etc.)". This form of disenchanted trust is "relational" (p. 25):

> Liberal democracies have emerged as a reaction of distrust to the traditional forms of power and authority such as monarchies and the church. As Mark Warren writes: 'More democracy has meant more oversight of and less trust in authorities'. Constructing a political arena in which people may confront their divergent interests and arguments means establishing a set of rules and procedures that allow a 'cold' yet guaranteed form of interaction not based on 'warmer' social relationships of trust.

The desire for liberal democracies to be inclusive has had a similar effect (p. 26): "Inclusiveness implies a transition from 'custom to code', because the more people are included within the same group, the less 'thick' relationships can be taken for granted."

More specifically, offline trust can be described as being based on a rational choice to trust (p. 29, italics removed):

> Here is an opportunity to enter a relation. If I choose this opportunity, I take some risks. I can cho[o]se to go on or to stay out, and I can base my decision on a rational calculus of the interests at stake or on an optimistic stance toward the willingness of the other party to honor

[2] All references to Origgi's work in this section are to Origgi (2013).

> my trusting him or her. But, in any case, I have a choice. I can choose to vote for this person because I think it is in his or her interest to take into account my interests (for example, because she or he wants to be re-elected in the future) but I can also choose to vote for someone else.

In Orrigi's view, online trust is very different from its offline relative. Rather than being based on a rational decision, "the form of trust that seems to reign over the Social Web is the most naïve and wild form of blind trust that we have ever experienced in mature societies" (p. 25). To wit, online trust is "epistemic trust, that is, trust in persons or systems through which we are able to extract relevant information" (p. 23), and the aim of Origgi's 2013 paper is precisely to "provide a definition of epistemic trust suitable to understand the massive trust we observe in social networks and discuss the complex relation between trust and information" (ibid.) bearing in mind that "[t]he social Web is first of all an epistemic engine that allows us to extract information about what happens around us. This use of the social web explains why it is so easy to trust in this virtual environment" (ibid.).

Thus, Origgi claims (p. 26), referring to Dwyer, Hiltz and Passerini (2007), that

> the disenchanted trust that defines our form of political participation doesn't seem to be the default attitude once we are on the Net. Social networking facilities have developed tremendously since 2007. Studies show that people develop online social networking even when the levels of reciprocal trust and comprehension of privacy and security issues are low. Most people who register on Facebook do not read the Terms of Service and, if asked, don't know whether they own or give away the information they make available on their profile pages.

Origgi expresses great amazement over how trusting people are online (p. 26):

> It is as if masses of reasonable individuals, who should be guided in their behavior according to mainstream views in social sciences – by maximization of interests and considerations of prudence and rationality, are willing to capitulate their judgement and responsibility of choice and join privately owned social networks and companies where they share personal information without the least clue of what these companies will do with these data, follow the first results of a Google search, confident that they will be brought to the relevant piece of content, base their

> judgements and evaluations on rankings produced by monopolistic companies. People seem willingly to throw away their privacy, their capacity of discrimination, their rights of choice and blindly defer to methods of filtering content and manag[ing] participation whose logic is deeply out of their control.

Indeed,

> [a] mixture of optimism, credulity and faith seems to be the dominant attitude that underlies the use of social networks, as if questions of privacy and security were not relevant to the development of this particular form of trust. Also, trust is a fundamental ingredient of social relationships, but it is unclear how people can trust millions of other users to make a fair use of the information they decide to share publicly. (p. 27)

How can this apparent discrepancy between online and offline interaction be explained?

> So, despite the blatant evidence of the risks on privacy, the control by private companies of most of the features and applications of the web, people seem to resist any form of diffidence or, at least of, prudence, and give away their personal data and relevant information *with a[n] unreasonable feeling of being part of a cooperative process of global democratization of means of expression*." (p. 27, italics added)

The suggestion is that one reason why people are so trusting online is that they mistakenly and irrationally believe they are part of a cooperative process when in fact this is not so:

> I don't think that the notion of relational trust I have tried to outline here says anything about our trust in information-dense environments such as the Internet and the Social Web. Trust in these environments is first of all a form of epistemic trust, that is, trust in persons or systems through which we are able to extract relevant information. (p. 29)

As for our use of search engines,

> [i]nformation on the Web is essentially social, that is, it depends on the pattern of social relations that informs the search algorithms about where to find it. We do not only develop social relations on Internet: we use social relations to extract knowledge from it. (pp. 29-30).

Hence,

> [s]earch-engines are thus powerful epistemic devices that we trust. We trust their capacity to perform the cognitive function of meta-memory (i.e. navigating through memory) in our place. Today, social networks allow other forms of hierarchisation of authorities, such as the 'follow' relation on Twitter, that ranks users' influence in terms of numbers of followers. (p. 30)

Yet, even though our trust on the web is taken to be generally blind there is, Origgi tells us, a sense in which we exert a limited form of vigilance when we use search engines like Google:

> Trusting a result of a search-engine like Google, or trusting information socially extracted from Twitter is a very different form of trust than the relational form I presented above. It is a form of epistemic trust that is based on different norms and heuristics of justification among which:
>
> - experience (I have double-checked information retrieved by Google and it was right, so I trust Google in the future),
>
> - our relation with epistemic authorities,
>
> - various reputational cues, on cognitive and epistemic properties of communication, and
>
> - structural features of social networks. (pp. 30-31)

The trust we have when we enter social networks, by contrast, is utterly blind and uncritical:

> The 'stance of trust' displayed in social communication through the Web is both fundamental and fragile: we need a minimal trust in order to get in contact with other people and enter new networks and conversations. That is why all the social networks that work well are characterized by the presence of trustful atmosphere, a 'halo' of trust and friendliness that is the trace that

> people share a stance of trust as a pre-conditions for constructing a communicative space. (p. 31)

What this illustrates, in Origgi's view, is a fundamental uncertainty in online interactions that cannot be completely eliminated on pain of losing out on the benefits of the web altogether: "As we can see, in all these cases, the trust we put at work is an epistemic attitude and not purely a matter of strategic choice: we cannot choose not to trust, because we would not have any alternative access to information" (p. 32) and so "a trustful attitude is a default condition to start any process of information seeking in information-dense environments". (p. 32).

Our situation on the web is therefore essentially asymmetric (p. 33): "the fact that we must share first in order to get something relevant from the social web puts us in an asymmetric situation that can creat[e] pernicious effects that explain the puzzle of trust I have started with."

Origgi's suggestion for how we can avoid the most naïve forms trust on the web is that we should combine a default trustful attitude with a strong reliance on reputational cues (p. 34):

> These two features of epistemic trust in web-based social networks [a default trustful attitude and strong reliance on reputational cues] may help solve the puzzle of the 'enchanted trust' on Internet in opposition with the 'disenchanted trust' typical of mature democracies and explain why people so naïvely trust on Internet. As I said, these two forms of trust are deeply different, the first one being a cognitive posture we need to take in order to filter the too thick amount of information we have to parse.

On a charitable interpretation, Origgi is here proposing not to "solve the puzzle" in the sense of showing that it was not one to being with, but rather to show how we can steer between trust and vigilance online in a way that avoids the most obvious negative effects of naïve reliance on online information.

3. Testable claims in Origgi's account of online trust

Our focus in what follows will be on Origgi's characterization of online trust. In particular, we will be concerned with the empirical basis of that characterization. Origgi's main claim, or hypothesis, is the following:

(H) Trust on the web (mainly in connection with search engines and online social networks) is blind, massive and naïve.

The support that Origgi mounts for this bold and far-reaching hypothesis consists of a number of empirical claims that are in principle testable. Regarding OSE, we have identified the following testable claims:

(OSE1) People follow the first results of a Google search.

(OSE2) People base their judgements and evaluations on rankings produced by private companies.

(OSE3) People using search engines have a feeling of being part of a cooperative process of global democratization of means of expression.

(OSE4) The kind of trust that is relevant to understanding use of search engines is not relational trust based on rational choice among cooperating agents.

(OSE5) Trust in search engines is an epistemic attitude and not (purely) a matter of strategic choice.

(OSE6) Trust in the results of a search engine involves assessments of experience ("I have double-checked information retrieved by Google and it was right, so I trust Google in the future"), our relation with epistemic authorities, various reputational cues, cognitive and epistemic properties of communication, and structural features of social networks.

Origgi also makes a number of empirical claims in connection with OSN, of which the most distinct ones are listed below:

(OSN1) People develop online social networking even when the levels of reciprocal trust and comprehension of privacy and security issues are low.

(OSN2) Most people who register on Facebook do not read the Terms of Service and, if asked, don't know whether they own or give away the information they make available on their profile pages.

(OSN3) People using social networks have a feeling of being part of a cooperative process of global democratization of means of expression.

(OSN4) The kind of trust that is relevant to understanding social networks is not relational trust based on rational choice among cooperating agents.

(OSN5) Online trust when using social networks is an epistemic attitude and not (purely) a matter of strategic choice.

(OSN6) Social networks (that work well) are characterized by the presence of trustful atmosphere, a 'halo' of trust and friendliness.

(OSN7) Users of social networks must share first in order to get something relevant from the social web.

Our strategy in the next sections will be to rate the extent to which these empirical claims are supported in relation to relevant recent empirical web research literature. Doing so presents some challenges, as some of Origgi's claims bear upon issues that current research on internet use does not in fact address or does address but only indirectly. For instance, there is little, if any, research that addresses the user's general and continued attitude towards the OSE and OSN. As we will see, most of the relevant literature concerns what users know about certain specific features, what motivates them to share or access certain specific content, etc. As a consequence, some of Origgi's empirical claims, though testable in principle, have not, to the best of our knowledge, actually been tested in practice. With this caveat, Origgi's claims can be nonetheless distributed into three categories: those whose truth-value is beyond reasonable doubt (whether true or false); those whose truth-value is either supported or undermined by the extant research, but where the matter has not been settled yet; and those whose truth-value is unclear at the time of writing either because there is important evidence both for and against or because evidence is missing. We will go through these categories, in that order.

4. Claims whose truth value can be established (whether true or false)

As we shall argue, at least two of Origgi's claims are true beyond reasonable doubt, namely OSE1 and OSE2, which are repeated here for ease of reference:

(OSE1) People follow the first results of a Google search.

(OSE2) People base their judgements and evaluations on rankings produced by private companies.

Moreover, we will also argue that only one of Origgi's claims is false beyond reasonable doubt, namely OSN7: users of social networks must share first in order to get something relevant from the social web. To be more exact, OSN7 turns out to be false on its most natural reading, while on another possible, but less natural, interpretation, it becomes almost trivially true. Let us now turn to the specifics, starting with OSE2.

There is no need to look into the empirical literature to substantiate the part of claim OSE2 concerning ownership of OSE. As we all know, the most popular online search engines, like Google Search and Bing, are operated by private companies (Google and Microsoft, respectively). Therefore, as long as users of OSE base their "judgment and evaluations" on the outcome of online search, these judgements and evaluations are indeed based "on rankings produced by private companies". While this part of OSE2 is a platitude, this does not mean of course that OSE2 as a whole is. Origgi writes in fact very little about the relation between these rankings, judgments and evaluations. What she says is compatible with the view that the basis for judgment and evaluation is exhausted by these rankings, but it is also compatible with the view this basis includes these ranking and perhaps also other information. Her writing is also consistent with the view that the basis for judgment and evaluation is context-dependent. The only possibility that she seems to exclude, is that users' own interests successfully screen off, as it were, the interests of private companies that are taken to be manifested in their rankings.

Yet, OSE2 remains sufficiently general to be counted as true, if the notion of something being a basis for judgment and evaluation is interpreted charitably enough, due to the fact that a number of studies

indicate that those rankings do in fact play a significant role in user's evaluations of the content they access online. Specifically, the empirical studies that directly support OSE1 – the claim that people follow the first results of a Google search – support OSE2 as well. It is well-established experimentally that users restrict their attention to top-ranked items of a Google search, a finding that has been confirmed by a variety of methods. A seminal study, reported in Pan *et al* (2007), used eye-tracking to determine whether users would privilege the ranking of a search item, or the relevance of the item relative to the initial search term. The study concluded that in case of conflict between relevance and ranking, the latter gets precedence. This result has been consistently replicated and is known under the label "top link heuristics" (Salmeron *et al*, 2013). This said, recent studies indicate that students follow more links to pages with unreliable information than to pages with reliable information, but spend more time on the latter (Rodicio, 2015); and that third-party interventions, such as a librarian suggesting reformulations of search queries, influence search behavior towards a greater responsiveness to relevance and reliability cues, and less towards ranking, as was shown in Leeder and Shah (2016). The latter study also supports conclusions drawn from observations in more natural settings, in particular those of Hargittai *et al* (2010). Nevertheless, it seems that OSE2 remains unchallenged as a thesis about the general non-student public lacking access to expert recommendations.

The last of Origgi's claim whose truth-value can be seen as established is OSN7, which we will argue is a false claim. OSN7 states, we recall, that users of social networks must share first in order to get something relevant from the social web. Our first remark is to note an ambiguity in the claim in question. OSN7 may be interpreted as stating either that that *every* user needs to share first in order to get something relevant from the social web, or that *some* users need to do so. Let us call the first interpretation "universal" and the second "existential", to reflect the difference in quantification over users. Now if OSN7 is read as a universal claim, it is factually false. The amount of information that one has to share online in order to create an OSN profile is actually fairly small – usually just a user name and an e-mail address. In return one can access all the information publicly accessible in the OSN. Non-trivial sharing of information is not necessary in order to access to access information

online, although it may severely limit the amount of suggestion that one receives from personalization algorithms and recommendation systems. Even so, it is not necessary for all to share in order for users to harvest the fruits of OSN.[3] If, on the other hand, Origgi's claim is taken, less naturally, in its existential sense, it is true, but only in a trivial sense. There cannot be massive access to relevant information, if no one is actually sharing. So, at least some users need to share in order for users to have access to information over an OSN. Equally obvious, the individuals who share need not do so to gain access to information relevant to *them*, for the same reasons that invalidate the universal interpretation of OSN7. Unfortunately for Origgi, once the necessity of an individual "entry price" or "blind buy-in" has been invalidated, her explanation of why individual users share (blindly or not) is left unsupported by empirical fact.

Let us dwell on the importance of OSN7 for Origgi's project. Being a philosopher (now working in France), Origgi may have been inspired, in her acceptance of OSN7, by the kind of reasoning that motivates Blaise Pascal's wager (Pascal, 1994/1669). Pascal argued, famously, that believing in god is an optimal choice. If god exists, you will be richly rewarded in afterlife; if not, then at least you will be endowed with a good Christian character. If, by contrast, you choose not to believe, then you will either be tormented in hell, if god exists, or have a less than fully admirable character, if he doesn't. In an equally famous essay, William James (1897) clarified the conditions of the wager by saying that the choice is "genuine" in the sense of being "living", "momentous" and "forced". It is living in the sense that both options – believing or not – are live ones, i.e. they are not already excluded but make some appeal, however small, to the person's beliefs. It is momentous because it is important: something of great value can be won and the opportunity is unique. Finally, the choice is forced because it is, in a sense, unavoidable: the utility of avoiding the choice by being agnostic is the same as the utility of not

[3] It has been argued on theoretical grounds that personalization algorithms and recommendation systems incur a risk of inducing "filter bubbles", echo chambers and confirmation biases (e.g. Sunstein, 2001, Pariser, 2011, Simpson, 2012), although the extent to which particular services do this is controversial (concerning Google, see Hannak *et al*, 2013). Hence, not sharing, or not sharing too much, may be a way to avoid these effects to the extent that they exist. But this is a normative issue, unlike Origgi's claim, which is descriptive in character.

believing. James made the general point that if a choice is genuine in this sense and there is no evidence to decide the matter, the best strategy is to "buy into" the claim in question, in Pascal's case the claim that god exists.

Similarly, Origgi seems to be suggesting that the decision to share information in a social network is a genuine choice in this sense. It is living because both options – sharing or not sharing – are live alternatives. It is, Origgi seems to think, momentous because it is important and the opportunity unique: information and social gratification can be won only if you share yourself. The choice whether to share or not to share is forced and unavoidable because we have to do one or the other. Finally, the choice is blind in the sense that we have to share in the absence of evidence that we get the information and social gratifications that we hope for, rather than, say, problems with integrity violations or cyber-bullying. It would now follow, by Jamesian reasoning, that sharing online is the right thing to do from a practical standpoint. So, we should blindly buy into Facebook or some other OSN and start sharing information with the other users.

This said, the wager is an argument for making a choice that is "blind" but, at least in one sense, not "naïve". The choice is blind because the recommended action represents taking a "leap of faith" in the absence of empirical information about the state of the world. It is not naïve because informed by decision-theoretic considerations. Origgi, by contrast, argues that online sharing is blind *and* naïve. Hence this reconstruction of her argument does not account for all the "phenomena".

Be that as it may, what we have pointed out, in effect, is that the choice whether to share is in fact *not* a genuine option in James's sense. It is not a genuine option because it is not momentous: a person wishing to benefit OSN membership does not have to share (non-trivial) information herself. It is sufficient that some users, or a significant number of them, share. It remains true, however, that a user may not be in a position to enjoy all the potential fruits of OSN membership unless she also shares herself.[4]

5. Empirically supported but not established claims

[4] Cf. our discussion of veteran users of OSN in section 6.

Let us now turn to those of Origgi's claims that, although their truth value cannot be established, are strongly supported by current web research. Falling into this category are OSN2, OSE5 and OSE6.

(OSN2) Most people who register on Facebook do not read the Terms of Service and, if asked, don't know whether they own or give away the information they make available on their profile pages.

(OSE5) Trust in search engines is an epistemic attitude and not (purely) a matter of strategic choice.

(OSE6) Trust in the results of a search engine involves assessments of experience, our relation with epistemic authorities, various reputational cues, on cognitive and epistemic properties of communication, and structural features of social networks.

Among them, OSN2 is the claim for which the most robust evidence is available. In a recent eye-tracking study, Steinfeld (2016) showed that most users who are not presented with a policy by default hardly ever click to read it, while those who do so merely skim through the text. Steinfeld also showed that default presentation increased users' knowledge of use of personal data, and encouraged reading with no relation to privacy attitudes. However, it is worth noting that OSN seldom present their policy by default. Steinfeld's study also provides an up-to-date review of the literature (as of Feb. 2016), which strongly indicates that it is practically impossible for users to read extensively privacy policies or to keep track of their changes. The first impossibility stems from the time it would take to read in relation to reasonable time constraint, and the second from the fact that many policies include the provider's right to change the policy without requiring user consent. Returning to OSE2, Facebook provides a link to its privacy policy but does not present it by default. Hence, if Steinfeld's results transfer from the controlled experimental setting to a more naturalistic one, it is quite plausible to subscribe to OSE2.

In addition, OSE5 and OSE6 are also well-supported by extant research, although the latter is not extensive. In fact, few studies have, as far as we can discover, actually looked at credibility assessment to online sources in a broad context of online information-seeking, rather than in the narrow context of a controlled experiments. The difficulties involved in addressing the issue in a natural setting are spelled out in (Hargittai *et al.*, 2010), which remains, as of today, one of the few studies in this

category. This study described the habits of "digital natives" (or "millennials") relative to information search, and their subsequent credibility judgments for results of online search. The study concluded, in particular, that even digital natives rarely use OSE without having first considered other sources to which they have extant trust relations, such as traditional gatekeepers (librarians and teachers) or peers (fellow students). Hargittai and her colleagues also showed that digital natives rely on their past experience and the reports of others' experience, as well as on 'reputational cues' such as house names (of brands and institutions), domains (such as ".edu" as opposed to ".com") and personal credentials of authors (such as diplomas, accreditation or endorsements).

Clearly the above findings support both OSE5 and OSE6. Hargittai *et al* (2010) also found that users often assess credibility based on a presumption of relevance that may lead them to misinterpret available cues. For example, some users trusted Google rankings to the point that they assumed that sponsored links were most relevant due to their placement higher in the search results, which is consistent with the hypothesis of a top link heuristics (cf. our discussion of OSE1 and OSE2 in section 3). Finally, studied users also displayed levels of digital literacy that varied considerably, which impacted their research skills. For instance, less than half of them were aware of the existence of Google Scholar, and less than one tenth used it (Hargittai *et al*, 2010). This lends support to Origgi's view that internet users rely on the same type of cues as they do with off-line epistemic authorities, but also lack in general the maturity to do it efficiently; at least this holds for the demographics studied by Hargittai and her colleagues.

Some studies have tried to replicate the natural approach of Hargittai *et al* (2010). Thus, Wineburg and Reisman (2015) looked at online sources and others, relative to disciplinary literacy of youth in historical sciences; Johnson et al. (2015) studied credibility assessment of information relative to health issues; and Kim and Yang (2015) looked at online information in the context of digital natives' civic engagement and political opinion formation. All these studies agree with the conclusion drawn by Hargittai and her colleagues, lending further support to OSE5 and OSE6.

It is here worth noticing that the support for OSE6 is stronger than that for OSE5. As for the former, the cues that Origgi hypothesize as the source of online trust match exactly those that Hargittai and her colleagues highlight in their empirical study.[5] For the purposes of evaluating OSE5, however, the support depends on how these cues are in fact used. The cues in question could very well enter into rational and strategic calculations akin to either maximization of expected utility or satisficing. In that case, the empirical findings would still support OSE6 but the support for OSE5 would be wholly or partially lost. OSE5 is fully supported by the extant research only insofar as the presence of reputational cues triggers non-strategic ("blind") trust.

6. Claims for which there is mixed or no evidence

In this section, we look at the claims that remain and in relation to which there is, in our opinion, no clear verdict coming from extant web research. The first claim that we will consider is OSN1: people develop online social networking even when the levels of reciprocal trust and comprehension of privacy and security issues are low. Origgi (2013, p. 26) refers the study by Dwyer et al (2007) in support of OSN1. Yet the evidence derivable from this study for the claim in question is somewhat mixed. For one thing, the study did not look at the comprehension of privacy issues, only at how they are perceived. In other words, there was no comparison of perception to reality. OSN1, by contrast, involves comprehension, not mere perception, of privacy and security issues. Moreover, at the time of the study, Facebook was not available to the general public and creating a Facebook profile required a verified association with a higher education institution, which is likely to have had a positive influence on trust and security, making the Facebook study largely irrelevant to the part of OSN1 which requires that reciprocal trust is low. Even so, the claim that people develop online social networking even when the level of reciprocal trust is low is supported by the part of the study that is about MySpace. In that connection, it was found that MySpace users develop online relations with unknown users more willingly than Facebook ones do, and MySpace did not require users to verify their identity.

Let us turn now to OSN4 and OSN5:

[5] There is no reference to Hargittai *et al* (2010), or any other empirical study of search engines, in Origgi (2013).

(OSN4) The kind of trust that is relevant to understanding social networks is not relational trust based on rational choice among cooperating agents.

(OSN5) Online trust when using social networks is an epistemic attitude and not (purely) a matter of strategic choice.

OSN4 can be undermined by arguing that the hypothesis that online trust is a form of relational trust, based on rational choice among cooperating agents, is methodologically fruitful and in line with observed data. Unfortunately for Origgi, there is plenty of evidence in favor of the relational model in the context of OSN. Extant Facebook research indicates that new users' strategies for sharing content do not respond to feedback from others. Rather, they seem to exemplify *social learning* in the sense of passive monitoring of others' strategies followed by imitation of those strategies (Wilson *et al*, 2012). These strategies are also known not to be negatively correlated with perceived risks for privacy (Burke et al., 2009). In other words, they persist in the face of such perceived risk. As for veteran users' sharing habits on Facebook, Twitter and YouTube, they do respond to feedback from other users, and they do so relative to motivations that fit well with a cost-benefit model, where risks involving failure of (relational) trust in other users and institutions are incurred in return for opportunities for social gratification (Lee and Ma, 2012). As of today, these results are robust enough to undermine OSN4: indeed, a cost-benefit analysis has all it takes, with regard to explanatory power, to account for online behavior of experienced users. While new users' habits are not well explained by this analysis, epistemic trust in the sense of Origgi does not explain their behavior either.

We will now turn to the assessment of OSN5. Some of the strongest evidence *against* this claim is provided by research that Origgi herself cites, namely Morris *et al* (2012), whose results have been recently replicated and extended by Jahng and Littau (2016). Both studies examine which features Twitter users rely on when assessing the credibility of messages whose authors are *outside* of their own networks (both offline and online). They showed that in these cases, users intentionally rely on features that are immediately accessible on a Twitter profile, namely the profile picture, the user name, and the capsule description. Furthermore, these features differ from the cues they agree that they

should use, like those that could be found by accessing a Twitter user's webpage. The reason is that the latter require an additional search effort to access, which is not compatible with the fast nature of information search on Twitter. Under those conditions, users unsurprisingly do not perform well relative to credibility assessments. But their poor performance can be blamed on the overall design of Twitter pages, and there are indications that the former could be improved by changing the latter (Jahng and Littau, 2016). Now, crucially, users seem to rely on these features to *assess a baseline of trust* prior to accepting the content of a tweet as factual. Hence, their attitude fits a cost-benefit model of relational trust, relying on cue-based heuristics. In the special case of Twitter, at least, such reliance fits within a cost-benefit model and thus undermines OSN5, showing that a non-relational model is far from indispensable in understanding online social networks and thereby counting also against OSN4.

Among the claims that are neither clearly supported nor clearly undermined by empirical web research we also find OSE4: the kind of trust that is relevant to understanding use of search engines is not relational trust based on rational choice among cooperating agents. There is, as far as we know, no direct evidence either for or against OSE4 which, like its cousin OSN4, is a claim on the theoretical or methodological level. As we argued, trust in the results of a search engine is compatible with a heuristics-based, cost-benefit assessment, where reputational cues would serve to fix a baseline (see our previous discussion of OSE6). While this observation to some extent undermines OSE4, there is no direct evidence against that claim, while there is, as we saw, direct evidence against OSN4 which is the corresponding claim for social networks. Hence, for all practical purposes, we have stalemate.

Then we have OSE3 and OSN3:

(OSE3) People using search engines have a feeling of being part of a cooperative process of global democratization of means of expression.

(OSN3) People using social networks have a feeling of being part of a cooperative process of global democratization of means of expression.

Neither of these claims has, to our knowledge, been tested *per se*. It might be the case that people do harbor the feelings in question, or that some do but other do not, or that no-one actually does. There is

simply no available data. In fact, OSE3 and OSN3 essentially amount to the assumption that the general public believes in the advertisement of the providers of search engines and social networks. Origgi may be inclined to believe that this is actually the case because she believes that user behavior is characterized by blind trust. From that perspective, OSE4 and OSN4 would count as arguments for OSE3 and OSN3, respectively, because the latter would be the "best explanation" of the former. However, we have seen that OSE4 and OSN4 are not actually supported by the research: the cost-benefit analysis that assumes relational trust is capable of organizing the data on online behavior. With the necessity of a non-relational model goes the support for OSE3 and OSN3 from inference to the best explanation. Then again, there is no concrete evidence against those two claims, and in all fairness, we can pronounce a stalemate regarding their validity as well.

Finally, there is OSN6 – the claim that social networks (that work well) are characterized by a halo of trust and friendliness – which is difficult to assess as it stands. No doubt online social networks use words such as "friends" and have "like" buttons, and there are no "enemies" and (at least at the time of writing) no "dislike" buttons. But OSN6 turns a blind eye to the phenomena of Facebook trolls, negative Twitter storms, negative comments on YouTube, and in general online drama. In fact, OSN6 seems to follow from the same implicit assumption that could explain acceptance of OSE3 and OSN3, namely that the general public buys into the superficial features of online networks. But this time the assumption is even harder to defend, for users would have to buy into these features to the point of becoming blind to whatever does not match them, and again there is no shortage of things that don't. It is true though that Facebook users can quickly deploy containment strategies by means of which unfriendly behavior can be neutralized, and there are several studies focusing on those strategies (for an overview, see Wilson *et al*, 2012). Hence, a weaker claim than OSN6 is probably defensible to the effect that users go to a great length to *maintain* a "halo of trust and friendliness", which is not the same as saying that the atmosphere of trust is the natural state of things.

7. Conclusion

Origgi's hypothesis is that trust on the web is blind, massive and naïve. According to her, this is so because online trust is epistemic rather than relational and cannot be explained by a cost-benefit analysis. According to Origgi, online trust requires a "blind buy-in", dictated by the presumption that one enters a cooperative communication situation in the absence of evidence of cooperation. On balance, however, we have questioned the relevance of epistemic trust in Origgi's sense in an account of online behavior relative to either OSE or OSN. As we have pointed out, extant research is not only compatible with a cost-benefit model based on relational trust – i.e. a model that makes trust neither blind nor naïve – but in a number of cases exploits that model fruitfully. The reasons for the massive sharing of information in the context of OSN may not be, as of today, well understood, but there are, in our view, few indications that cost-benefit models based on relational trust are inadequate.

This said, we also found much to agree with in Origgi's account from an empirical perspective. She is definitely right in her claims about users' strong reliance on private companies such as Google for ranking search results. Also, her claims about users' ignorance about integrity issues are plausible, although we noted that there is some additional complexity that needs to be taken into account in a full analysis. Finally, Origgi is on the whole right in her remarks about various cues that people use in their assessment of search results, in addition to the ranking provided by the search engine itself.

Still, Origgi's central contention that users of social networks have to blindly buy into Facebook or other social networks in an act resembling a "leap of faith" seems to us not to be correct. Users must not share (non-trivial) information first in order to get something relevant from the social web, such as information and social gratification. What is required is that some users, or a significant number of them, share; not that all do. It remains true, though, that a user may not be in a position to enjoy the full potential benefits of social network interaction unless she is willing to share information herself.

While Origgi's account of online trust is only partially supported by empirical research and there is in some cases evidence to the contrary, it is surely true that many people are too trusting online. Origgi's recommendation that we should further digital literacy is also recommended by internet scholars, whether the focus is on search engines or social networks. Thus, we fully agree with Origgi's

assessment that current lack of online alertness "should be balanced by a vigilant attitude of epistemic responsibility not only from the perspective of the producers of information, but also from the perspective of consumers of information" (Origgi, 2013, p.36, italics removed).

Finally, we would like to make it perfectly clear that we greatly appreciate Origgi's boldness in devising new far-reaching hypotheses about the nature of internet interaction. In doing so, she manifests the great relevance of epistemology and philosophy generally as providing possible ways of interpreting empirical web research, which is often scattered and fragmentary. Moreover, the distinctions between various forms of trust introduced by Origgi raise a number of important questions, e.g. concerning how they relate to each other and how they fit, if at all, into a decision-theoretic framework. In this article, we could only address some of these exciting issues, leaving others for another occasion. No doubt, Origgi's conceptual work will be seen to be indispensable in any future endeavor to make sense of interaction on the web.

References

Burke, M., Marlow, C., and Lento, T. (2009). Feed me: motivating newcomer contribution in social network sites. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 945–954.

Cha, M., Benevenuto, F., Haddadi, H., and Gummadi, K. (2012). The world of connections and information flow in twitter. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 42(4): 991–998.

Dwyer, C., Hiltz, S., and Passerini, K. (2007). Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace. AMCIS 2007 Proceedings, 339-.

Grice, H. P. (1989). Logic and conversation. In Studies in theWay ofWords, pages 22–40. Harvard University Press. First published in 1975. (* NOT REFERRED TO *)

Hannak, A. et al. (2013). Measuring personalization of web search, Proceedings of the 22nd international conference on World Wide Web.

Hargittai, E., Fullerton, L., Menchen-Trevino, E., and Thomas, K. Y. (2010). Trust online: Young adults' evaluation of web content. International Journal of Communication 4(27): 468-494.

Jahng, M. R. and Littau, J. (2016). Interacting Is Believing Interactivity, Social Cue, and Perceptions of Journalistic Credibility on Twitter. Journalism & Mass Communication Quarterly 93(1): 38-58. First published online in 2015.

James, W. (1897). The will to believe, reprinted 1956 in The will to believe and other essays in popular philosophy. New York: Dover, 1-31

Johnson, F., Rowley, J., and Sbaffi, L. (2015). Modelling trust formation in health information contexts. Journal of Information Science 41(4): 415-429.

Kim, E.-m. and Yang, S. (2015). Internet literacy and digital natives civic engagement: Internet skill literacy or Internet information literacy? Journal of Youth Studies: 1–19.

Lee, C. S. and Ma, L. (2012). News sharing in social media: The effect of gratifications and prior experience. Computers in Human Behavior 28(2): 331 – 339.

Leeder, C., and C. Shah. (2016). Measuring the Effect of Virtual Librarian Intervention on Student Online Search. The Journal of Academic Librarianship 42(1): 2-7.

Morris, M. R., Counts, S., Roseway, A., Hoff, A., and Schwarz, J. (2012). Tweeting is believing?: understanding microblog credibility perceptions. In Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work, ACM, 441–450.

Origgi, G. (2004). Is trust an epistemological notion? Episteme, 1(01): 61–72.

Origgi, G. (2005). What does it mean to trust in epistemic authority? In 7th Annual Roundtable of Philosophy of Social Science, pages 1–26, New York. Barnard College, Columbia University., HAL archives-ouvertes.fr. Accessed October 29, 2015, from https://hal.inria.fr/file/index/docid/, 59465/filename/Roundtable_of_Philosophy_of_ Social_Science.pdf.

Origgi, G. (2012). Epistemic injustice and epistemic trust. Social Epistemology 26(2): 221–235.

Origgi, G. (2013). Democracy and trust in the age of the Social Web. Teoria Politica 3: 23–38.

Pan, B., H. Hembrooke, T. Joachims, L. Lorigo, G. Gay, and L. Granka. (2007). In google we trust: Users' decisions on rank, position, and relevance. Journal of Computer-Mediated Communication 12(3): 801-823. http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2007.00351.x/full

Pascal, B. (1964). Penseés, ed. Léon Brunschvicg. Paris: Garnier. First published in 1669.

Rodicio, H. G. (2015). Students' evaluation strategies in a Web research task: Are they sensitive to relevance and reliability? Journal of Computing in Higher Education 27(2): 134-157.

Salmerón, L., Kammerer, Y., & García-Carrión, P. (2013). Searching the Web for conflicting topics: Page and user factors. Computers in Human Behavior 29: 2161–2171.

Simon, H. A. (1959). Theories of Decision-Making in Economics and Behavioral Science. The American Economic Review 49(3): 253–283.

Simpson, T. W. (2012). Evaluating Google as an epistemic tool. Metaphilosophy 43(4): 426-445.

Sperber, D., Cl´ement, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., and Wilson, D. (2010). Epistemic vigilance. Mind & Language 25(4): 359–393.

Sperber, D. and Wilson, D. (1995). Relevance: Communication and cognition. Oxford: Blackwell.

Steinfeld, N. (2016). "I agree to the terms and conditions": (How) do users read privacy policies online? An eye-tracking experiment. Computers in Human Behavior 55: 992-1000.

Sunstein, C. R. (2001) Republic.com. Princeton University Press.

Wilson, R. E., Gosling, S. D., and Graham, L. T. (2012). A review of Facebook research in the social sciences. Perspectives on psychological science 7(3): 203–220.

Wineburg, S. and Reisman, A. (2015). Disciplinary Literacy in History. Journal of Adolescent & Adult Literacy 58(8): 636–639.