



# The Cost of Truth. Motivations of a Pragmatist Trust-Conditional Approach to News Evaluation

Dario Compagno

## ► To cite this version:

Dario Compagno. The Cost of Truth. Motivations of a Pragmatist Trust-Conditional Approach to News Evaluation. Versus: Quaderni di Studi Semiotici, 2018, 127, pp.275-290. 10.14649/91356 . hal-01949981

**HAL Id: hal-01949981**

**<https://hal.science/hal-01949981>**

Submitted on 9 Mar 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DARIO COMPAGNO

# The Cost of Truth

## Motivations of a Pragmatist Trust-Conditional Approach to News Evaluation

This paper argues that fake news is a consequence of some flaws of the media system and suggests a direction to deal with this phenomenon. We begin by observing that the traditional media are regulated by two competing systems of values, truthfulness and proximity, and that the second is factually more important than the first. We argue that social media has maximised proximity and hence, has completed the separation of trust and truth. We propose to call “trust-conditional evaluation” the model of truth under which the social media operates today. We show that such a model eventually leads to idealistic and relativistic consequences, unless it is constrained by a pragmatic attitude. We suggest then a pragmatic constraint: Statements can be said to be true only relatively to an evaluation effort, therefore the truth value of a statement corresponds to the cost of falsifying it. We conclude by discussing some potential implementations in the design of algorithms using trust ties in social networks in order to evaluate statements.

**Keywords:** trust-conditions; fake news; negative realism; cost function; fairness.

### 1. Introduction

How come it is so easy to demonstrate *in vitro* that some piece of journalistic information is a fake and at the same time, so difficult to prevent its massive circulation (Keersmaecker and Roets 2017)? Answers appealing to generalised human stupidity (Ferraris 2017) appear to at least hide some aspects of the problem. The same goes for those who pretend that fake news is just bullshit, that is, that it is indifferent about truth (Dieguez 2018). A more complex answer, questioning the way in which we naively apply the concept of truth to the news, is potentially capable of generating insights about how to deal with fake news and perhaps about the nature of truth itself. In fact, the social media displays our inherent hermeneutical limitations, shattering the collective belief of a realist *ante rem* truth independent from *in rebus* interpretation and trust.

## 2. News is not Truth

We may say that the news system, as it has evolved within Western societies, is flawed. Its flaws are determined by a contradiction between two different sets of values. On the one hand, we have a completely ideal imperative that is supposed to motivate the social existence of journalism, but that actually plays just a minor role in it. The news is supposed to *truthfully* represent the human society and its goings-on. The end of journalism as a truthful reflecting mirror is then said to drive towards the end of democracy. In fact, some argue, the difficulty of finding a coherent standpoint for observing the development of society weakens its contractual existence (Rouvroy and Stiegler 2015).

A second set of values at work in journalism has been identified by sociologists (Galtung and Ruge 1965). These researchers were interested in identifying the actual, and not the ideal, criteria determining which events deserve to be published. They found a set of general criteria, called the criteria of newsworthiness, that can be used to explain which statements are published by newspapers at a certain time (see Harcup and O'Neill 2016 for a recent assessment). These criteria can be roughly summarised by one general principle: To interest a public, news has to be close to it. Such criterion of *proximity* states that to sell the same number of copies, a hurricane on the front page has to kill many more foreigners than fellow citizens. The readers' disposition to consume information depends on their subjective standpoint and the information supply adapts to this demand.

Significantly, truth is not amongst the criteria of newsworthiness. In a sense, sociologists present truth as an implicit trait not worth investigating. The point of Galtung and Ruge was to understand why some presumably true stories are covered by the media, while other equally true stories are not. Still, perhaps beyond the authors' original intentions, they found an autonomous explication for information propagation, parallel to the truth verification procedures performed by the media. Let us concentrate on this hypothesis according to which a long-lived conflict between two different sets of values (truthfulness and proximity) may be the cause for today's generalised exposure to false information.

There is no need to entirely deny any connection between news and truth. Most evidently, the way in which we talk about the media, does often refer to their truth value. Even the law imposes a certain respect of truth determined by juridical means. For example, depending upon national regulations, false information about a person or an organisation may need to be rectified because it is considered defamatory. It is interesting to notice, however, that not all false information face the same consequences, and that even for the law, some other criteria take the upper hand over truth. In fact, even a true statement may be considered defamatory if its publi-

cation is not justified by public interest. The right to know and the right to inform, in reference to which the spread of information is regulated in Western societies, relate only obliquely to the truth value of what is written.

Such a mismatch between news and truth is also visible whenever we consider how we judge and class newspapers. Tabloids cannot be held responsible for false statements in the same way high-profile journals are. *The New York Times'* journalists follow some stricter practices, complying with the commitment not to publish anything that has not been verified thoroughly, or not to exaggerate the content of an article to the point it distorts their honest-to-goodness interpretation of the facts. Other publications are instead even expected by readers to offer uncertain speculations or ideological reasonings. The choice of a newspaper by readers is rooted in cognitive consonance: One expects a *vraisemblable* in which the true is seen under the aspect of the likely. It does not help that human beings are subject to a confirmation bias (see Nickerson 1998). Psychologists found that we tend to look for and remember only facts and arguments confirming our hypotheses, while we systematically avoid critical ones. As a consequence, our personal interests and preferences influence the truths we have access to.

The point is that truth has a social weight often underestimated by logicians. One cannot believe in whatever he or she wants, because beliefs depend on and contribute to determine a person's positioning in the social world. Researchers who have studied the ties of trust linking individuals to one another found a connection between what is believed to be *true* and the social connections in which people *trust*; for example, changing one's belief demands to consider how this affects other people's reactions (Lewicki and Bunker 1995). In this perspective, truth is not at all independent from trust, that is, to evaluate a statement demands also to look beyond the world as it "really" is.

In particular, trust in a newspaper entails a certain belief of truthfulness, and even if newspapers have since always been accused of lying, the news system managed until recently to induce a collective belief of trustable truthfulness. Readers cannot verify individually every piece of information, but collectively they were able to converge on a consensual judgement about at least the most important social facts (such as about who won the last national elections). Sociologists have talked since decades of the social construction of reality: Facts need to be collectively endorsed to become true (Berger and Luckmann 1966). Actual assignations of truth values to published sentences depend on social practices that are something more than the simple apprehension of facts.

The evidence we are quickly collecting from sparse domains of knowledge appears to show that journalism has tolerated and made use of truth much more than it has founded its practices on it. Proximity, trust and truth

are more connected than we would like to think<sup>1</sup>. The modern *régime de vérité* was grounded on a regulated access to the public discourse based on trust in newspapers. Today, the Internet frees this access and as a consequence makes it harder to converge collectively on a shared core of truths.

### 3. Social Media Maximise Proximity

We have referred to truthfulness and proximity as two contrasting sets of values associated to news production and diffusion. We could resume these preliminary considerations by saying that truth is what *ought* ideally to be published, while news is what *gets* actually published, and news is based on proximity. Let us now observe why social media completes the detachment between news and truth.

Social media today occupies an important role in news distribution, as more and more people access journalistic contents through social media accounts (Mercier and Pignard-Cheynel 2018). Information production and reception are therefore affected by the logics of the social media, mediating supply and demand. It is important to remark then that the social media works by maximising proximity: its algorithms provide a higher chance to be read to contents that are close to the user. Newspapers have found in the social media a perfect ally, because their working logics are ultimately identical. Better still, the social media provides a measurably effective way to bring media contents to readers who appreciate them. As a consequence, the “natural” propension of the media system to aim at proximity (§ 1) is “artificially” amplified by the social media, to the further expense of truthfulness.

At the beginning of its development, the algorithm for sorting web pages created by Google (called PageRank) did not work in this way. Google calculated the importance of a page by evaluating its network of citations. A page referred to by many pages was considered to be more important than one that was not; also, to be referred to by an important page gave more importance than to be referred to by a less important one. This recursive calculation managed to converge onto metastable values for all the pages taken into consideration by the algorithm. This means that, for Google, importance was still an *objective* feature of web pages: It depended on the full network of pages referring to each other, while the user performing a query was not influential. Therefore, the same query performed by two different users produced exactly the same result. Today, Google’s algorithm is supposed to not work in this way anymore because it has adopted more subjective ways of content evaluation, taking the user into account.

<sup>1</sup> Proximity is not equivalent to trust. Still, in this paper we suggest that proximity and trust were entangled in the traditional media, more so with social media.

Proximity is calculated by social media on the basis of the user's history of interactions with other users and with contents. Facebook's algorithm (formerly named EdgeRank) introduced this idea: What matters for the user of a social network is not the impersonal, objective importance based on other contents, but its *subjective* importance, that is, what other users close to him or her in the network find important. In Facebook, importance is calculated *ad hoc* for each couple user–content. As a consequence, anything may be important for somebody, and nothing is important once and for all. The user's activity becomes an explicit variable in the calculation, aimed at improving the adherence between his or her expectations and the sorted results.

It should be noted that neither Google's PageRank nor Facebook's EdgeRank are interested in truthfulness. What matters is only importance, expressed by hyperlinks or captured more subtly by recording traces of the user's interest (comments, “likes”, clicks). As distributors of content, current social algorithms are blind to truth. If some false information was referred to by many important pages, Google considered it important. Similarly, if some false information is appreciated by many who are close to the user in the social network, Facebook considers it important. It is up to the user to critically determine whether a certain piece of content is also true. And it appears that, removed from the institutional clues and guardrails surrounding traditional news production, users alone are often incapable of performing this task (Keersmaecker and Roets 2017).

In this way, social algorithms have completed the removal of truth from proximity, and shown that these two values, simultaneously and ambiguously endorsed by the traditional media, are actually far removed from each other. The strength of the social media is to inductively build a model of the society's structure from traces of individual activities. To this end, what gets published and shared (news) is most useful, while what ought to be published and shared (truth) is a bias to be cornered.

Filtered by the social media, the collective belief of trustable truthfulness in the traditional media appears to be questioned, as peer-to-peer trust ties replace institutional ones. The information supply is decomposed and then, recomposed around the users, building *ad hoc* worlds based on their ties of trust (Mercier and Pignard-Cheynel 2018). As a matter of fact, this produces barriers in information propagation: News do not circulate uniformly and do not reach the farthest points of the network, that is, the people who are less likely to believe them. Researchers also found that retractions never reach all those who read some fake news (Vosoughi, Roy and Aral 2018). Fake news exists in close, although possibly quite large, communities shielded from larger consensual evaluation; even debunking has been proven to have limited effect (Zollo *et al.* 2017). The point is then to understand how we can fluidify the social propagation of news.

To do this, we need a general model of truth capable of grounding the development of new algorithms of information management. But let us first detail the model of truth currently at work in social media.

#### 4. Trust-Conditional Anti-Realism

A *realist* view of truth, seeing it as independent from the practices of production and interpretation of discourse (Glanzberg 2016), is today at work in virtually all discussions about fake news. Common sense and experts alike resort to a realist correspondence theory, according to which, mainly: Each statement has to be either true or false, its evaluation depends on the knowledge of a fact, and its correct truth value can be obtained with a commensurate effort. As an example, «Donald Trump is the President of the United States of America» must be either true or false, the truth value of this statement depends on whether it is an “unquoted” fact that Donald Trump occupies that role at the present time, and such knowledge can be obtained by journalistic investigation.

Our hypothesis is that fake news is above all a scandal for this naively shared perspective on truth. Discussants of fake news appear to accept the idea that some piece of information, considered to be fake, is evidently false *because* there is an objective way to determine what is true. We will argue later, in a pragmatist perspective (§ 4), that the preceding utterance is incorrect in this form, as we can only determine what is false and never what is true. For the moment, let us take a closer look at the contrast between, on the one hand, the realist takes on truth grounding many discussions about fake news and, on the other hand, the anti-realist take fuelled by the doing of contemporary social media.

Realist analytic philosophers defend the idea according to which the variety of uses of the words “true” and “false”, observable in linguistic practices, stem from an original immutable concept of truth. The question is, what notion of interpretation follows from realism? If truth values are always defined (logical binarism) and depend directly and univocally on extra-linguistic truthmakers (facts), then (correct) interpretation should ideally *preserve* this assignation. Interpretation is needed to commensurate language and facts, but correct interpretation is perfectly transparent. Whenever interpretation reshapes reality, this is by definition a vicious activity, removed from the ideal of a unique, honest mirroring of facts. In this perspective, information technology should limit itself to preserving the pristine human access to reality.

*Anti-realist* theories of truth, such as Michael Dummett’s (1978), instead reverses the hierarchy of interpretation and truth. Truth is considered to be



the end of many language games, the reason we talk, but it cannot easily be defined in general terms. Each language game is “played” differently, allowing for different “moves” that are in fact different assertion and verification procedures. As “winning conditions” of language games change from game to game, we can only vaguely grasp what “winning” means in general. In this perspective, it is not at all inconceivable that truth may *evolve* along the refinement of social practices, which introduce new assertion and verification procedures. It could then be heuristic to think about truth as the contingent result of technological development in information management. In this view, truth is a concept whose understanding should be derived from the study of linguistic practices, including journalistic ones, as there is nothing before and above some family-resemblances among real-world uses of the word “true”.

It should be added that realist theories are often built on a *correspondent* concept of truth: A statement (truth-bearer) is true if a certain fact outside of language and interpretation (truth-maker) exists. Facts are exactly the same for all people and all languages, constituting an objective benchmark for evaluation. Anti-realist theories instead usually endorse a *coherence* concept of truth: A statement is true if it is coherent with a whole (of actual or acceptable statements, of propositions, of beliefs). This makes possible a plurality of truth systems, each starting from different premises or adopting different rules to elicit new knowledge.

With something more than a pun, we may say that for realists, statements have truth conditions, while for anti-realists, statements have *trust conditions*. By this, we mean that in a realist perspective, evaluation stands just on the way in which the world is. And, for realists, to understand a statement is to grasp the way in which the factual world has to be in order for the statement to be true (truth-conditional semantics). By contrast, in an anti-realist perspective, if we aim to derive the truth value of a statement starting from other statements, we need to begin with some premises. In a real-world scenario, accepting some premises corresponds to trust someone of their validity. Therefore, to look at the world is not sufficient to evaluate statements; one also needs to look at the social ties structuring the production and interpretation of discourse.

The main idea at the basis of what one may call *trust-conditional evaluation* is that truthfulness depends on trustworthiness<sup>2</sup>. We might even imagine a *trust-conditional semantics* stating that to understand a statement is to grasp the way in which the social world has to be in order for the statement to

<sup>2</sup> Two research threads are related to this idea. On the one hand, truth-conditional pragmatics as proposed by Recanati (2010) and, on the other hand, computational approaches to trust as in Huang and Nicol (2010) or Jøsang, Marsh and Pope (2006).



be trusted. We argued that truth is what should ideally be said and published, while news is what actually gets said and published. Accordingly, truth-conditional evaluation may work for identifying the ideal conditions under which statements are true (from a godly or bird's eye view, so to speak), but trust-conditional evaluation may explain better what happens in real world language production and interpretation. In fact, trust conditions cannot tell whether a statement is true, but *for whom* it is true.

We claim that the algorithms of proximity developed by the social media (§ 2) are the eventual elementary implementation of an anti-realist perspective: Something is true if my close peers say it is. Whenever users want to verify a news item, they usually resort to search for more information, for more statements capable of supporting the first one. They stay online and go look for more sources; the practice of going offline, “out there”, to perform perceptual judgements so to directly verify what they read has decreased by degrees. This is obvious, given that we have become interested in what happens on a much larger scale than we used to earlier in the past. It is probably not a coincidence that the criteria of newsworthiness (§ 1) were elaborated by sociologists working on international news, farther from the reader's possibility of direct verification.

Now, errors can be tolerated by society, as it was in modern journalism, if it does not obstruct the functioning of the system, namely the production of a trustable truthfulness. But some claim that we are currently shifting towards more distributed forms of trust, that is, towards social practices that are less based on indirect confidence in traditional institutions and more on direct confidence in individuals, made possible by peer-to-peer technology (Botsman 2017). This leans towards a scenario in which numerous individual judgements count more than institutional ones, impeding the collective convergence of beliefs.

As a reaction, we may therefore wish to pursue one of three alternative aims: Either we look for ways to reprimatinate the realist idea of truth (perhaps by reinvigorating top-down trust in the institutions), or we are happy about the idealist dissolution of a unique reality of facts, or we search for a third model of truth capable of working in a distributed, “headless” system of trust. We will now make some considerations for this third research scenario.

## **5. A Pragmatist and Realist Take on Truth**

Since George Boole, all along the evolution of modern logic, truth has been seen as the result of calculation, building up from the semantics of elementary logical components to full statements' truth values. Arguments by anti-realist philosophers have shown that the more restrictive limita-

tions we give to evaluation, *the harder the calculus* of truth values is. As an example, intuitionist natural deduction systems, in which the bivalence principle does not hold, cannot produce derivations as easily as the ones based on classical logic (Pelletier 1999). In a very precise sense, anti-realist views of truth are harder to take into account practically, but their results are stronger, as what can be proved by anti-realist (intuitionist) reasoning can also be proved by the classical realist one, but not vice versa. As a consequence, anti-realism *per se* should not be thought as an easier way to truth, as if anyone could believe in whatever he or she preferred. Still, anti-realism often brings about idealistic and relativistic consequences: If truth is a matter of coherence and deductibility, different sets of premises will lead to different realities. Aiming at a global consensus, we then need to guide the choice of premises and of reasoning rules; we need some guardrails making computation *even harder* so as to assure the convergence of opinions.

We have argued above (§ 2) that the social media performs a calculation of proximity that takes into account the user as a variable. Contents are sorted on the basis of their adequacy to different users' histories, and there is no universal sorting. We may then say that social media exclusively calculates in terms of elementary trustworthiness (§ 3). This is a harder calculus than plain first-order logic evaluation based on realist formal semantics, arbitrarily assigning univocal truth values to elementary logical components. However, pure trustworthiness falls easily into relativism, and this is what seems to happen nowadays with fake news. Let us recall that trust conditions do not tell whether a statement is true or false, but for whom it is true, and so trust conditions alone do not permit to build a bottom-up consensual social reality. Instead, we want to find a way to tell whether a statement *is* true or false working consensus up from trust, without recurring to "realist" shortcuts (i.e., assuming that truth *must* be independent from interpretation). The solution we propose is to look at how pragmatists define truth so as to perfect the anti-realist trust-conditional approach described above, with the aim of making it "pragmatically realist" instead than "naively realist" (James 1907).

According to Charles S. Peirce's view, very briefly, truth is what is obtained as the result of research in the very long run (Peirce 1877). Peirce did believe in a real world independent from interpretation, capable of a definitive assignation of truth values to statements, but the apprehension of this real world demands *infinite time*. Scientific knowledge, in the meanwhile, is fallible because it can be revised at any time. As a consequence, truth should not be thought as an actual present reference but as an ideal final referring. According to Peirce, this was enough to make him a realist; according to today's standards, his position makes him more of an idealist. In particular, as for Peirce (1868) direct intuition of external objects is

impossible and knowledge always depends on other knowledge; his take on truth is more easily describable in terms of coherence than in terms of correspondence<sup>3</sup>.

If attempted to be seen from a Peircean perspective regarding ways to improve how social media works, truth would be the final convergence of an algorithm based on trust. Whenever *the entire community* converges on a judgement, that becomes truth. In this paper, we want to elaborate on this idea, that we consider fundamentally correct, by making it implementable by algorithms. It is easy to see why Peirce's idealistic assumption cannot work as it is in a real-world scenario. We need a reachable reality, an algorithm needs to converge in finite time; we cannot wait and see whether or not the entire community agrees. Therefore, if truth cannot be practically obtained, because at best it would take too much time, then we should at least find a perspective from which we can manage *falsehood*.

These considerations lead to another pragmatist perspective, namely that of Umberto Eco (1997; 2007). Eco accepts most of Peirce's doctrine but modifies it slightly to avoid the fall into idealism. Very briefly, the infinite progression of research is stopped by the intuitive component of perceptual judgements (called by Eco "primary iconism") anchoring them to reality *in finite time*. To see is to see reality instantly, even if through the varied lenses of culture. Now, for Eco there can only be *negative correspondence* between statements and facts, that is, one can say when a statement does not correspond to reality, but one can never tell when it does correspond. So, perception somehow goes beyond pure coherence without replacing it. Experience answers with "noes" to some of the questions human beings ask. Now, if truth were a binary, as in traditional realist theories, from a "no" given by experience, one should easily be able to deduce a "yes", so to affirm the negation of the rejected statement (logical binarism), but that is not the case. The "noes" of experience do not result in positive knowledge. In logical terms, for Eco, falsehood is a *semidecidable* set: One may find direct proof of a statement's falsity but not of its truth. This makes Eco an anomalous realist, but still a realist to all practical extents. We could compare Eco's approach with what Montaigne wrote about lying:

<sup>3</sup> Peirce's realism could be taken to a stronger degree, especially in reference to his later work. Peirce's thought could be interpreted as if what we take as true at any given time is so taken because we represent it as what would be taken as true if investigation were protracted far enough. An idealistic asymptotic limit cannot be removed from Peirce ("far enough"), but it could be seen as a way of practically dealing with knowledge instead than as a way of asserting the non-detachability of reality from knowledge. In our reading, instead, knowledge depends on signs and semiosis cannot conclude as it is constitutively open to further interpretation. As a consequence, Peirce's reality, being *definitive*, becomes also ultimately unknowable, therefore interfering with his larger doctrine. Umberto Eco's solution to this conundrum is presented below.

If a lie, like truth, had only one face we could be on better terms, for certainty would be the reverse of what the liar said. But the reverse side of truth has a hundred thousand shapes and no defined limits (Montaigne, *Essays*, I: 9, "On Liars").

From knowing that something is a lie we cannot derive a truth, as if true and false were the two sides of a coin. Montaigne still believed that truth existed in univocal and definite form; for Eco, in real life, we only deal with the "reverse sides of truth" and reality is a *pure formless constraint* whose nature we can only interpolate from the "thousand shapes" of tentative false statements. The unquoted facts imagined by "realists" are replaced by constraints that can never be quoted, as they can be seen only in negative by observing the actual efforts and developments of human epistemic activity. In this light, it is correct to say that *no news is true, but some news is provably false*<sup>4</sup>.

It should be clear why Eco's approach is viable for rethinking truth in a pragmatist and realist way that is also, at least in principle, implementable by algorithms. If verification demands infinite time, falsification instead can converge more easily, as it is directly grounded on perceptual judgements. Now, on a large scale, it is practically not possible to check who has or has not falsified a statement with his or her own point of view, but we can assume that people routinely do convince their peers that they have. William James wrote about this distributed correspondence grounding coherence:

Truth lives, in fact, for the most part on a credit system. Our thoughts and beliefs 'pass,' so long as nothing challenges them, just as bank-notes pass so long as nobody refuses them. But this all points to direct face-to-face verifications somewhere, without which the fabric of truth collapses like a financial system with no cash-basis whatever. You accept my verification of one thing, I yours of another. We trade on each other's truth. But beliefs verified concretely by somebody are the posts of the whole superstructure (James 1907: 80).

In fact, if we opposed correspondence and coherence, as two mutually exclusive ways to reach truth, we would have to deal with a dangerous dichotomy. Who in the social body should be given the right to access truth directly? Certified journalists, for example, should be trained to "see for themselves"? We are not convinced that consensus should result from top-down censorship, giving the right to talk only to an institutional selection of qualified experts. We are also not convinced that individual readers may be taught to manage large amounts of information better without computational support. We suggest instead to look for practical implementations of negative constraints to assure the global convergence

<sup>4</sup> Eco's approach recalls Karl Popper's asymmetry of verification and falsification, John Austin's analysis of the term "real" and Jacques Lacan's *réel*; a detailed discussion goes beyond the scope of this paper.

of a trust-conditional social algorithm. Accordingly, the bottom-up trustable truthfulness we are looking for, becomes what individuals are allowed to believe given the negative constraints of what society collectively knows should not be believed.

## 6. The Cost of Truth

In conclusion, we are going to sketch how the pragmatist and realist approach presented above (§ 4) may lead to an improvement in the way social media works. Reality is a resistance to interpretation (Eco 1997), and truth is a way to express and possibly quantify this resistance. A truly pragmatist take on truth “in so far forth” (James 1907) is an evaluation of statements making our best efforts. Evaluation in this sense is an activity, that as such comes with a cost. And the higher the quality of truth reached, the higher the cost to falsify it. In plain terms, *the truth value of a statement is the cost to falsify it*. Hence, we need algorithms to compute this cost.

The social media already estimates trust through proximity, now the point is to make them estimate truth from trust. This proposition aims to take into account the social dimension of truth, without making it arbitrary. This could extend to ground evaluation on empirical measures rather than on logical atoms. The calculus we have in mind is closer to network analysis than to propositional logic. It could be implemented in many ways, and here we only suggest a theoretical frame. The following are some principles upon which a pragmatist algorithm of truth could rely:

- 1) A statement is true relatively, not to a fact but to an effort of falsification (at a first glance, all statements are true, extreme effort falsifies any statement).

- 2) The human effort needed to falsify a statement depends on a complex cost (cognitive, social, technological, economical and other conditions determine how hard it is to falsify a statement in a certain community at a certain time).

- 3) The cost of human falsification can be modelled by the computational cost of an algorithm processing exclusively the trust conditions in a social network (it is possible to guess how hard it is for someone to falsify a statement by looking at his or her ties of trust).

- 4) Therefore, a statement’s truth value can be estimated by a cost function as global trustability score (from individual trust conditions formalizing for whom a statement is locally true, a function can derive how much the actual society would have to change for that statement to be considered globally true).

What happens today with fake news is that groups of people adopt local beliefs in closed social environments in which internal trust ties are

too strong. They just “choose to believe” in something. For example, stating that the Earth is round is very difficult to falsify, because this statement is coherent with a large number of other accepted ones (it is “analytical” in the sense given to this term by W.V.O. Quine), therefore a person rejecting this belief should as a consequence break a significant number of trust ties he or she entertains with more reasonable persons. But today, the social media does nothing to contrast the diffusion of statements with a very low cost of falsification. An indicator of resistance to our will to believe could instead help control the spread of fake news, by informing about how unlikely a statement is from a global perspective, that is, about how much the society would have to “pay” to accept a certain statement as true.

The assumption grounding our proposition is that a fully open network of individual evaluations is sufficient to obtain in finite time a convergent evaluation if statements are weighed according to their social cost of falsification. Social media platforms need to implement *opinion inertia*, so to speak, in order to convert trust into truth. Returning to our example, claims about a flat Earth circulate too freely on the social media, because their dissemination is not regulated by a global evaluation of their social cost. A better algorithm is needed: Independently from the fact that all of my closest friends believe that the Earth is flat, the diffusion of such statements should be mitigated by an indication of how other communities react to it. Something is true if also others, who are at a distance, are reluctant to falsify it. On the contrary, if the social network’s global constraints are not dealt with, one will always find someone so remote from common sense (i.e., from the “center” of the network) that the person might be convinced to believe in anything one says. In our view, local proximity has to be counterbalanced by a global criterion of *farness*, and not by a reference to facts as they “really are”.

Global constraints are what the social media currently lacks and the reason they have a loose grab onto reality. Constraints could be implemented in many ways: By accompanying news with a trustability score, therefore still leaving to individuals the task of evaluation, or more strongly, by limiting the circulation of low-scoring statements. Scores could be calculated automatically, starting from the users’ traces or could result from explicitly distributed evaluation. What is most important is that, here we are not promoting a conservative system in which innovation and deviancy are impossible or are censured by the many, but a system in which innovation and deviancy have meaning exactly because they are shaped by the many and have an impact upon them.

## 7. Conclusion

The future of the car is not the horse. It is unlikely that as a species we will renounce information technology because of its drawbacks. It is also unlikely that humans will become better at cognitively managing large amounts of information about subjects remote from their direct experience and competence, in a world becoming increasingly complex and rich with sophisticated technologies. More likely, interpretation will be assisted by algorithms. Fake news is not an essential by-product of algorithms, we just need better algorithms to help individuals. We argue that trust conditions alone carry the risk of relativism. However, if they are coupled with a cost function, modelling the social cost needed to reshape the entire community's ties of trust in accepting some new statement, they can be used to help the social body to converge on a global consensus.

**Dario Compagno**

Université de Lorraine

Ile du Saulcy

Metz (France)

dario.compagno@univ-lorraine.fr

## References

BERGER, P.L. and LUCKMANN, T.

1966 *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*, Garden City, Anchor Books.

BOTSMAN, R.

2017 *Who Can You Trust? How Technology Brought Us Together and Why It Might Drive Us Apart*, New York, PublicAffairs.

DE KEERSMAECKER, J. and ROETS, A.

2017 “Fake news’: Incorrect, but hard to correct. The role of cognitive ability on the impact of false information on social impressions”, *Intelligence*, 65, pp. 107-110.

DIEGUEZ, S.

2018 *Total Bullshit! Aux sources de la post-vérité*, Paris, PUF.

DUMMETT, M.

1978 *Truth and Other Enigmas*, Cambridge (MA) and London, Harvard University Press.

ECO, U.

1997 *Kant e l'ornitorinco*, Milano, Bompiani.

2007 *Dall'albero al labirinto*, Milano, Bompiani.



FERRARIS, M.

2017 *Post-verità e altri enigmi*, Bologna, Il Mulino.

GALTUNG, J. and RUGE, H.M.

1965 “The Structure of Foreign News. The Presentation of the Congo, Cuba and Cyprus Crises in Four Norwegian Newspapers”, *Journal of Peace Research*, 2, pp. 64-91.

GLANZBERG, M.

2016 “Truth”, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Fall 2018 Edition.

HARCUP, T. and O’NEILL, D.

2016 “What is News?”, *Journalism Studies*, 18, 12, pp. 1470-1488.

HUANG, J. and NICOL, D.

2010 “A Formal-Semantics-Based Calculus of Trust”, *IEEE Internet Computing*, 14, 5, pp. 38-46.

JAMES, W.

1995 [1907] *Pragmatism*, New York, Dover.

JØSANG, A., MARSH, S. and POPE, S.

2006 “Exploring Different Types of Trust Propagation”, in K. Stølen, W.H. Winsborough, F. Martinelli and F. Massacci (eds.), *Trust Management*, Berlin and Heidelberg, Springer, pp. 93-104.

LEWICKI, R.J. and BUNKER, B.B.

1995 “Trust in relationships: A model of development and decline”, in B.B. Bunker and J.Z. Rubin (eds.), *Conflict, cooperation, and justice*, San Francisco, Jossey-Bass, pp. 133-173.

MERCIER, A. and PIGNARD-CHEYNEL, N.

2018 *#info. Commenter et partager l’actualité sur Twitter et Facebook*, Paris, Editions de la Fondation Maison des sciences de l’homme.

NICKERSON, R.S.

1998 “Confirmation bias: A ubiquitous phenomenon in many guises”, *Review of General Psychology*, 2, 2, pp. 175-220.

PEIRCE, C.S.

1992 [1868] “Some Consequences of Four Incapacities”, in *The Essential Peirce*, Vol. 1, Indianapolis, Indiana University Press.

1992 [1877] “The Fixation of Belief”, in *The Essential Peirce*, Vol. 1, Indianapolis, Indiana University Press.

PELLETIER, F.J.

1999 “A History of Natural Deduction and Elementary Logic Textbooks”, in J. Woods and B. Brown (eds.), *Logical Consequence: Rival Approaches*, Oxford, Hermes Science, pp. 105-138.

RECANATI, F.

2010 *Truth-Conditional Pragmatics*, Oxford, Oxford University Press.

ROUVROY, A. and STIEGLER, B.

2015 “Le régime de vérité algorithmique”, *Socio*, 4, pp. 113-140.

VOSOUGHI, S., ROY, D. and ARAL, S.

2018 “The Spread of True and False News Online”, *Science*, 359, pp. 1146-1151.

ZOLLO, F., BESSI, A., DEL VICARIO, M., SCALA A., CALDARELLI G., SHEKHTMAN, L., HAVLIN, S. and QUATTROCIOCHI, W.

2017 “Debunking in a world of tribes”, *PLoS ONE*, 12, 7, online id: e0181821.