



# LUND UNIVERSITY

## Knowledge, truth, and bullshit: Reflections on Frankfurt

Olsson, Erik J

*Published in:*  
Midwest Studies in Philosophy

*DOI:*  
[10.1111/j.1475-4975.2008.00167.x](https://doi.org/10.1111/j.1475-4975.2008.00167.x)

2008

[Link to publication](#)

*Citation for published version (APA):*

Olsson, E. J. (2008). Knowledge, truth, and bullshit: Reflections on Frankfurt. *Midwest Studies in Philosophy*, 32, 94-110. <https://doi.org/10.1111/j.1475-4975.2008.00167.x>

*Total number of authors:*  
1

### General rights

Unless other specific re-use rights are stated the following general rights apply:  
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

# Knowledge, Truth, and Bullshit: Reflections on Frankfurt

Erik J. Olsson  
Lund University

Olsson, E. J. (2008). Knowledge, truth, and bullshit: Reflections on Frankfurt. *Midwest Studies in Philosophy*, 32, 94-110. Blackwell.

**Abstract:** This paper addresses two aspects of Harry G. Frankfurt's work on truth and what he calls "bullshit" – roughly, talk unconnected to the truth. In his short book *On Truth*, Frankfurt argues that truth is important for essentially instrumental reasons because of its usefulness as a basis for successful action. My main aim in the first part of this paper is to explore the connection between Frankfurt's work on the value of truth and recent work by myself, carried out partly in collaboration with Alvin I. Goldman, on the value of reliabilist knowledge. The second part concerns a puzzle raised by Frankfurt in connection with his discussion of the nature of bullshit. Frankfurt believes that a society can prosper only if there is shared agreement on a great many truths. He also thinks, however, that our society is affected by an "immeasurable flood of bullshit" prohibiting the truth from being generally known. And yet, our society does prosper. I explore various ways of solving the puzzle, drawing partly on recent work in social epistemology.

## 1. Frankfurt's *Meno* challenge

In his book *On Truth* (OT), Frankfurt argues – against postmodernism and other relativist philosophies – that there is an objective distinction to be drawn between the true and the false, and that the truth plays an essential role in our lives. Even those who present themselves as denying the tenability of the distinction must agree that this denial is a position that they *truly* endorse (OT, 9). They must agree that the statement expressing their rejection of the distinction truly and accurately describes their attitude towards it, which makes the relativistic stance ultimately incoherent.

We need, then, to acknowledge as genuine a distinction between what is true and what is false. But why, exactly, is truth so important to us? Frankfurt's basic answer amounts to an appeal to the instrumental value of truth in our daily affairs. It is surely extremely helpful to know the truth about what to eat and what not to eat, how to raise our children, where to live, and many other mundane matters (OT, 34-35). Truth, then, "often possesses very considerable

practical utility” (OT, 15). On a larger scale, engineers building, say, a bridge rely heavily on true assessments of various important parameters, such as the durability of construction materials (OT, 22). Moreover, they must ascertain, with reliable accuracy, both the obstacles that are inherent to the implementation of construction plans and the resources that are available for coping with those obstacles. Successful engineering depends therefore quite obviously on its practitioners being closely in touch with an objective reality that is independent of any particular point of view. The same could be said of architects and physicians whom we wouldn’t consult unless we believed them capable of forming objectively true statements in their respective areas of expertise. There is in all these contexts a clear difference between getting things right and getting them wrong and therefore a clear difference between true and false.

Although the level of subjectivity is greater when it comes to historical analyses and to social commentaries, there are even in these fields important limits to the interpretations that can reasonably be imposed on the phenomena. “There is”, as Frankfurt puts it, “a dimension of reality into which even the boldest – or the laziest – indulgence of subjectivity cannot dare to intrude.” (OT, 24).

Truth, Frankfurt observes, is important even in moral matters. This is so even if we concede that evaluative judgments themselves lack truth value. The reason is that whether we endorse such a judgment or not will depend on factual statements that do have truth values. Thus we may come to endorse an evaluative judgment of a person’s moral character on the basis of statements describing that person’s behavior in concrete cases. Those statements can be true or false, and they need to be true in order for our endorsement to be reasonable. Generally, we take things to be good or bad because of certain beliefs we have about those things. Thus we may hold one thing to be good because we believe it will increase our wealth or make us happier. If, upon closer examination, those beliefs turn out to be false, we tend to withdraw our initial positive sentiment towards the thing in question.

Because of the great practical usefulness of truth, we even have reason – Frankfurt thinks – to “love” it. Love is here construed, following Spinoza, as “joy with the accompanying idea of an external cause” (OT, 39), where by “joy” is meant “that passion by which the ... [individual] passes to a greater perfection” (OT, 41). If, in other words, a person experiencing joy in this sense identifies an object externally causing this passion, then, Spinoza believes, the person is rightly said to love that object. Now, truth, Spinoza and Frankfurt agree, is indispensable in enabling us to stay alive, to understand ourselves, and to live fully in accord with our nature (OT, 47). Once we recognize this, we must therefore love truth. Frankfurt

concludes that “[p]ractically all of us do love truth, whether or not we are aware that we do” (OT, 48).<sup>1</sup>

Frankfurt’s discussion of the value of truth does not end here, but I believe nonetheless that this short summary contains the most important components of his view. It also hints at what epistemologists will perceive as a shortcoming of Frankfurt’s presentation: the failure to distinguish clearly between the value of *truth* and the value of *knowing* the truth. Frankfurt moves effortlessly from the one to the other. Thus, he observes that people “need to know the truth about what to eat and what not to eat” etc. (OT, 35), an observation that is used to support the claim that they “require truths to negotiate their way effectively through the thicket of hazards and opportunities that all people invariably confront in going about their daily lives” (OT, 34-35).

Once the common distinction is drawn between truth in general, i.e., true belief, and truth that qualifies as knowledge, the question arises whether it is the former or the latter that has instrumental significance. Is true belief in general instrumentally valuable, even if the belief in question falls short of knowledge, or is it knowledge only that has practical worth? The answer, I submit, is that true belief in general has instrumental value. It is not necessary for a true belief to be valuable in this sense that it qualifies as knowledge.

Suppose, to take Plato’s example in his dialogue *Meno*, that you wish to embark on a journey to Larissa without having any idea of what direction to take. Now you encounter a trustworthy-looking, but in fact unreliable, guide who happens to give you correct information on the matter. Having listened to her, you form a true belief concerning the location of Larissa. Surely you are better off now, practically speaking, than you were before you consulted the guide. It is instrumentally better for you to possess true information as to the location of Larissa even if in fact you do not know that information to be correct.

Similarly, our engineer is better off having a true belief concerning the robustness of a given building material than having no or false information, even if her true belief fails to qualify as knowledge. For suppose that our engineer is handed a measurement instrument which happens to be somewhat unreliable but which nonetheless gives a correct reading in this particular case. Surely, this will leave our engineer in a better position, practically speaking, than she was before the measurements were conducted. It is instrumentally better for her to possess true information as to the robustness of the material than to be ignorant or in

---

<sup>1</sup> Frankfurt also believes that deception is quite common, especially in the form of bullshit, of which “[e]ach of us contributes his share” (*On Bullshit*, 1). But how can this be if all of us indeed love truth? Frankfurt’s answer, I presume, is that we are not (always) aware of our love for truth. Alternatively, one may speculate, truth may not be the only thing we love, and sometimes when our interests collide the truth is sacrificed. I return to Frankfurt’s discussion of bullshit in the second part of this essay.

error, even if she cannot be said to know that this information is true. It seems correct to say, then, that true belief in general is instrumentally valuable, not just true belief that is known to be true. We have reason, then, to love truth in general, not just to love knowledge, at least if we accept the Spinoza-Frankfurt definition of love.

But these remarks only serve to raise another, considerably more delicate and contentious, issue: given that true belief in general is valuable, how can it be shown that knowledge is even more precious? Ever since Plato, philosophers have thought that knowledge represents the most perfect grasp of reality that a person can ever hope to attain. Knowledge is more valuable than things which fall short of knowledge. Something can fall short of knowledge in many different ways, e.g. because it is simply false or not firmly held. Moreover, a true belief can fail to qualify as knowledge due to the less than reliable way in which it was acquired. Maybe Frankfurt need not provide a solution to the so-called *Meno* problem – the problem of accounting for the distinct value of knowledge – for his limited purposes of defending the importance of truth to a more general audience. But epistemologically more sophisticated readers are likely to remain unsatisfied with the present state of affairs.

What should a solution to the *Meno* problem look like in order to be congruent with Frankfurt's theory of the value of truth? One possibility would be to find a value other than instrumental value that knowledge has in greater degree than true belief in general. From a systematic perspective, however, it would be more satisfactory to learn that knowledge has more of the *same* value. Hence, from Frankfurt's point of view the first alternative to consider is certainly the option that knowledge is *instrumentally* more valuable than true belief in general.

## 2. Reliabilist solutions

Although there has been a lot of work on the *Meno* problem in recent years,<sup>2</sup> the idea that knowledge is more valuable than true belief in an *instrumental* sense is not one that currently enjoys great popularity. Indeed, Jonathan Kvanvig rejects it already in the first chapter of his influential 2003 book *The Value of Knowledge and the Pursuit of Understanding*.<sup>3</sup>

Why are philosophers skeptical to the idea of the extra value of knowledge being instrumental in kind? Many have been persuaded by the proposed counterexamples that surface in the literature, starting with Plato and his Larissa example. The latter is intended to show that whatever surplus value knowledge may have which true belief in general lacks, that

---

<sup>2</sup> See for example Jones (1997), Swinburne (1999), Riggs (2002), Zagzebski (2003), Kvanvig (2003) and Sosa (2003). For a recent overview, see Pritchard (2007).

<sup>3</sup> For a critical discussion, see Olsson (2007).

value cannot be instrumental. For there is, Plato claims, no practical difference between knowing the way to Larissa and merely having a true belief to that effect. In both cases, you are likely to reach your destination, the likelihood being also the same.

A further obstacle to giving a Frankfurt style solution to the *Meno* problem is that, while Frankfurt is reasonably explicit when it comes to the nature of truth and of bullshit, he never clarifies his use of the term knowledge. There are, nonetheless, reasons to believe that he would not object to a reliabilist construal of that concept. At one point, he stresses the importance for engineering that technological assessments are made with “reliable accuracy” (OT, 22), and later he shows appreciation of having a view that is not only true but also “reliably grounded in the relevant facts” (OT, 77). These quotations, to be sure, do not strictly speaking commit Frankfurt to a reliabilist definition of knowledge, but they do show that reliability is a notion to which he attaches some importance. In the following, knowledge will be taken in the (process) reliabilist sense of “true belief acquired through a reliable process”.<sup>4</sup>

Unfortunately, counterexamples similar in spirit to the Larissa example have been leveled directly against the reliabilist construal of knowledge. Thus, Linda Zagzebski (2003) has argued that just as a good cup of espresso does not get any better because it was reliably produced, so too a true belief does not become more valuable due to the fact that it was obtained through a reliable process. Examples of this kind have convinced most epistemologists working on the *Meno* problem that reliabilist knowledge is no more valuable, instrumentally or otherwise, than true belief in general.

Nevertheless, there are strong reasons to believe that epistemologists have generally overestimated the force of Zagzebski’s espresso analogy. It is a correct observation that, if the goodness of the espresso has been confirmed, the further information that it was reliably produced doesn’t make it taste any better. By the same token, if a belief is already assumed true, the further information that it was reliably produced doesn’t make that belief “more true”. But this argument assumes that the only virtue of reliable acquisition is that such acquisition indicates the quality (flavor or truth, as the case may be), of the thing produced. As we shall now see, this assumption is simply false.<sup>5 6</sup>

---

<sup>4</sup> Epistemologists will normally insist on an anti-Gettier clause in any definition of knowledge. Since the Gettier problem plays no role in this paper, I have chosen to simplify matters by not including an anti-Gettier clause in the analysis of knowledge. Classic formulations of the reliabilist position can be found in Goldman (1976) and (1986).

<sup>5</sup> For a critical discussion of the espresso analogy, see Olsson (2007).

<sup>6</sup> The account of the value of reliabilist knowledge that follows was first proposed in Goldman and Olsson (to appear). In that paper, which was written in 2006, it is referred to as “the conditional probability solution” and figures as one of two rather different suggestions for how to come to grips with the *Meno* problem. Parts of the present approach can be found in different places in the literature, e.g. in Armstrong (1973) in his reply to an

If a reliable coffee machine produces good espresso for you today, it can normally produce a good espresso for you tomorrow. The reason is that the coffee machine will normally be at your disposal tomorrow too; and, being reliable, it will probably produce one more good espresso. By the same token, if you have acquired a true belief that the road to the left leads to Larissa, then usually this same method will be available to you the next time around; and, being reliable, it will once again guide you correctly to your destination. There is more to reliable acquisition than meets the eye. The fact that a thing was reliably produced does not only indicate the quality of that thing; it also indicates that more things of the same quality can be produced in the future.

Reliabilist knowledge will tend to multiply. Having it increases the probability of getting more of the same. Unreliably acquired true belief, by contrast, does not share this feature to the same extent. If your true belief that Larissa is to the right was acquired through an in fact unreliable method, there will be an increased tendency for that same method, being unreliable, to lead, upon reemployment, to a false belief. Hence, the probability of your attaining further true beliefs is greater conditional on your possessing knowledge than conditional on your possessing a true belief in general.

Obviously, the extent to which the possession of knowledge raises the probability of future true beliefs depends on a number of empirical regularities. One is that people rarely face unique problems. Once you encounter a problem of a certain type, you are likely to encounter a problem of the same type at some later point. Problems that arise just once in a lifetime are few in number. If you are going to Larissa, the question of what is the best turn to take will probably occur more than once. Moreover, if a particular method solves a problem once, this same method is usually available to you the next time around. If you have an on-board GPS computer, you can use the navigation system to solve the problem of what road to take at the first crossroads. Unless the GPS system is stolen shortly thereafter – in most neighborhoods a rare event – this same method is also available to you when the same question is raised at the next crossroads. A further empirical fact is that, if you have used a given method before and the result has been unobjectionable, you are likely to use it again, if available, on similar occasions. Having invoked the navigation system once without any apparent problems, you have no reason to believe that it shouldn't work again. Hence, you decide to rely on it also at the second crossroads. Finally, if a given method is reliable in one situation, it is likely to be reliable in other similar situations as well. I will refer to these four empirical conditions as *non-uniqueness*, *cross-temporal access*, *learning* and *generality*, respectively.

---

objection raised by Deutscher. Williamson (2000), 100-102, presents a similar theory focusing exclusively on the special case of beliefs with temporally related contents.

To see even more clearly what roles these regularities play, suppose S knows that p. By the reliabilist definition of knowledge, there is a reliable method M that was invoked by S so as to produce S's belief that p. By non-uniqueness, it is likely that the same type of problem will arise again for S in the future. By cross-temporal access, the method M is likely to be available to S when this happens. By the learning assumption, S is likely to make use of M again on that occasion. By generality, M is likely to be reliable for solving that similar future problem as well. Since M is reliable, this new application of M is likely to result in a new true belief. Thus the fact that S has knowledge on a given occasion makes it to some extent likely that S will acquire further true beliefs in the future. The degree to which S's knowledge has this value depends on how likely it is that this will happen. This, in turn, depends on the degree to which the assumptions of non-uniqueness, cross-temporal access, learning and generality are satisfied in a given case.

Clearly, no corresponding conclusion is forthcoming for unreliably produced true belief. While non-uniqueness and cross-temporal access are usually satisfied quite independently of whether or not the method used is reliable, there is less reason to believe that an unreliable method that yields a correct belief on its first occasion of use will also yield a correct belief on the second occasion. This blocks the step from the availability of the method on the second occasion to the likely production of true belief on that occasion.

If we combine this account with Frankfurt's observation that truth has instrumental value, it follows that knowledge is even more valuable. For a state of knowledge has instrumental value in virtue of being a state of true belief. In addition, knowledge makes further states of true belief (and, indeed, knowledge) more likely. Knowledge is, in that sense, instrumentally more valuable than true belief in general.

My second point is that knowledge is particularly stable.<sup>7</sup> A true belief that qualifies as knowledge will tend to stay put and will not so easily go away. As Williamson (2000) notes, stability of true belief is of practical importance in the carrying out of complex actions over time. Frankfurt's example of engineers planning and building a bridge illustrates the point. It is crucial that the assessments upon which the engineering decisions are based will not change in the construction process. The thesis that true beliefs that are stable promotes successful action to a greater degree than true beliefs in general – what I call the Stability Action Thesis (SAT) – is firmly supported by pre-systematic judgment and will not be further argued for. However, it also needs to be established that reliable acquisition promotes stability. I call this

---

<sup>7</sup> The stability theory was proposed in Olsson (2007).



the Reliability Stability Thesis (RST). Together RST and SAT imply the Reliability Action Thesis (RAT): reliabilist knowledge promotes successful practical action over time.

It remains to justify RST, the thesis that reliable acquisition of true belief is conducive to stability. The main part of the justification amounts to showing that, if one is using an actually unreliable method to acquire a given belief, the unreliability will tend to be detected in due time. Once the method has proven to be unreliable, beliefs that were acquired through that method will tend to be withdrawn. By contrast, the chance that doubt will be shed on an actually reliable process is lower, and it is therefore less likely that beliefs arrived at through such a method are later discarded.

In order to make this likely, appeal will be made again to some empirical background assumptions. We will need all the assumptions that were invoked in the previous argument: non-uniqueness, cross-temporal access, learning and generality. It will be assumed, in addition, that, while our inquirers may sometimes succumb to wishful thinking and other less reliable paths to belief, most of their belief acquisition processes are in fact reliable. This will be expressed by saying that they are *overall reliable*. According to the assumption of *basic reliability*, the inquirer has at her disposal some method that is basic in the sense that it can be used to resolve conflicting verdicts. Visual perception is a good example of a method that is basic in this sense: in cases of dissonant beliefs, we can often resolve the issue by stepping forward and taking a closer look. Furthermore, inquirers will be supposed to be *track-keepers* in the sense that they keep a record of where they got their beliefs from. According to a further assumption, inquirers view their beliefs as *corrigible*, meaning that they typically don't stick to their beliefs no matter what. More precisely, an inquirer who finds a given belief false is likely to question the reliability of the method by means of which that belief was formed. Moreover, once a given belief-acquisition method is classified as dubious by the inquirer, all beliefs that were obtained solely or mainly through the use of that method are also, to some extent, in doubt. Together these conditions express a sense in which an inquirer's cognitive faculties are in good order.<sup>8</sup>

Why, then, should there be a tendency for reliably acquired (true) beliefs to be stable? Equivalently, why should there be a tendency for unreliably acquired (true) beliefs to be discarded? Consider the following sequence of events:

---

<sup>8</sup> Cf. Williamson (2000), p. 79. Maybe some of these assumptions can be weakened. I would welcome an inquiry into this matter, but I do not intend to pursue such an inquiry myself at this point. It suffices for my present purposes to show that, given some reasonably realistic assumptions, true beliefs that are reliably acquired will tend to be more stable than true beliefs in general.

1. S acquires a true belief that p through method M in response to problem P.
2. S faces a problem of the same type as P in the future (*non-uniqueness*)
3. S still has access to method M (*cross-temporal access*)
4. S uses M again (*learning*)
5. S now acquires a false belief, say that q
6. S becomes aware of a conflict between q and the verdict of one of her reliable methods (*overall reliability*)
7. The falsity of q is confirmed by some basic reliable method, such as visual perception at short distance (*basic reliability*)
8. S gives up her belief that q and considers M to be unreliable (*corrigibility*)
9. S notes that her (true) belief that p was also acquired through M (*track-keeping*)
10. S gives up her (true) belief that p (*corrigibility*)

Given the assumption of generality of reliable methods, stating that a method that is reliable now will probably continue to be so in the future, this unfortunate sequence of events is more likely, if M is unreliable, than it is, if M is reliable. Why? Because step 5 – the acquisition of a false belief – is more likely if M is unreliable. The other steps are, we may reasonably assume, equally likely for reliable and unreliable methods. Combining this result with our previous observation that a true belief that is stable over time will tend to be more useful in action, we may conclude that (reliabilist) knowledge is instrumentally more valuable than a true belief in general.

Let us apply this abstract piece of reasoning to Frankfurt's engineering example. Suppose that our engineer is confronted with the problem of assessing the robustness of a given piece of building material. Her method for assessing the robustness is using a certain measurement device. Using this device, she comes to the conclusion that the degree of robustness equals x, for some number x. The next day, the engineer faces the similar problem of assessing the robustness of some other material. Having the measurement device still at her disposal, she decides to use it again. Unbeknownst to the engineer, however, the device misrepresents the robustness of the new material, yielding a robustness value y much higher than the actual value. As a result, the engineer forms a false belief to the effect that the robustness of the new material equals y when in fact it is much lower. The construction work now proceeds on the basis of an overestimated robustness value for the second material which is used, so our story goes, in one of the critical parts of a new bridge. A few days later, that part of the bridge breaks down before our engineer's very eyes. She now becomes aware of a conflict between

her previous measurement and the confirmed fragility of the material leading to her giving up her belief in the robustness of the second material and reconsidering the reliability of the measurement device. It now occurs to our engineer that she applied the same device also on the first material. Accordingly, she concludes that this measurement cannot either be taken at face value, although it is in fact correct. The claim was that an episode like this one, terminating in the loss of a true belief, is more likely to unfold if the method used is unreliable as opposed to reliable.

I believe that there is a further, albeit more indirect, way in which these approaches to the *Meno* problem are congenial to Frankfurt's thinking on the value of truth. Frankfurt insists, as we saw, that truth is instrumentally valuable and that we should therefore be truth-seekers as well as truth-tellers. Yet he is, of course, well aware that in special circumstances it might be better not to know or tell the truth. To take one of his own examples, "a lie may divert us from embarking upon a course of action that we find tempting but that would in fact lead to our doing ourselves more harm than good" (OT, 75-76). To be sure, it would have been better if we could have been thus diverted without any recourse to lying, but it would be unrealistic to suppose this always to be possible. The course of action having the greatest practical value all things considered may be one that involves not telling the truth. Does this mean that truth is, after all, an overrated concept?

No. What it means, Frankfurt must agree, is that all that can reasonably be said is that truth is *normally* instrumentally valuable. That truth is normally instrumentally valuable means that the inference from "This is true" to "This is valuable" is a defeasible one in the following sense: we may infer the latter from the former, provided that we do not possess further information to the effect that the circumstances are special in ways that would undermine the conclusion. The situation is exactly analogous when it comes to the value of knowledge. Knowledge is distinctively valuable value *in all normal cases*. There is a defeasible inference to be drawn from "This is knowledge" to "This is distinctively valuable". The conclusion can be withdrawn should new evidence emerge suggesting the violation of one or more of the empirical regularities previously alluded to. The situation should be considered relevantly special if it was a one-shot case, so that the reliable method could not be reemployed, or if the inquirer considers herself incorrigible, to take just two examples.

### **3. Frankfurt's puzzle about bullshit**

Acknowledging an objective distinction between truth and falsity, Frankfurt believes, as we saw, that having true beliefs is important mainly for practical reasons. This is so not only for individual inquirers, but also for society as a whole:

Civilizations have *never* gotten along healthily, and *cannot* get along healthily, without large quantities of *reliable factual information*. They also cannot flourish if they are beset with troublesome infections for *mistaken* beliefs. To establish and to sustain an advanced culture, we need to avoid being debilitated either by error or by ignorance. We need to know – and, of course we must also understand how to make productive use of – a great many truths. (OT, 34-35)

The use of the first person plural indicates that Frankfurt is thinking of knowledge that is shared by the members of the society.

Yet there is also deception, one form of which is the central concern of Frankfurt's *On Bullshit* (OB). In the book, Frankfurt contrasts bullshit not only with truth-telling but also with lying, one of his conclusions being that the truth-teller and the liar are more closely related than either is to the bullshitter. The reason is that, while both the truth-teller and the liar react, in their own different ways, to the truth, the bullshitter is generally unconcerned with the way things really are:

Someone who lies and someone who tells the truth are playing on opposite sides, so to speak, in the same game. Each responds to the facts as he understands them, although the response of the one is guided by the authority of the truth, while the response of the other defies that authority and refuses to meet its demands. The bullshitter ignores these demands altogether. He does not reject the authority of truth, as the liar does, and oppose himself to it. He pays no attention to it at all. (OB, 60-61)

Indeed, being thus unconcerned with the truth is, in Frankfurt's view, the very essence of bullshit (OB, 33-34).

What is worst, then, bullshitting or lying? Frankfurt thinks that the former is more seriously harmful, at least in the long run:

Both in lying and in telling the truth people are guided by their beliefs concerning the way things are. These guide them as they endeavor either to describe the world correctly or to

describe it deceitfully. For this reason, telling lies does not tend to unfit a person for telling the truth in the same way that bullshitting tends to. Through excessive indulgence in the latter activity, which involves making assertions without paying attention to anything except what it suits one to say, a person's normal habit of attending to the ways things are may become attenuated or lost. (OB, 60)

The bullshitter, Frankfurt speculates, gradually becomes a victim of her own humbug: by allowing her reporting to be systematically influenced by ulterior motives she gradually becomes unreliable in her private believing, which makes her unfit to tell the truth. At any rate, the activity of lying, by contrast, is conceptually impossible in the absence of a genuine concern for the truth. By definition, a person cannot lie without believing that there is a truth that is elsewhere to be found. There is no reason to believe, therefore, that lying should corrupt its practitioner's epistemic character in a way similar to how bullshitting allegedly does. For this reason, "bullshit is a greater enemy of truth than lies are" (OB, 61).

For essentially the same reason why bullshitting is harmful to the individual, "bullshitting constitutes a more insidious threat than lying does to the conduct of civilized society." (OT, 4-5). Consequently, "whatever benefits and rewards it may sometimes be possible to attain by bullshitting, by dissembling, or through sheer mendacity, societies cannot afford to tolerate anyone or anything that fosters a slovenly indifference to the distinction between true and false" (OT, 33). For "[i]f people were generally dishonest and untrustworthy, the very possibility of productive social life would be threatened." (OT, 69)

Now we cannot help observing – Frankfurt thinks – that our society happens to be one in which deception, in all its shapes and forms, is extremely common:

After all, the amount of lying and misrepresentation of all kinds that actually goes on in the world (of which the immeasurable flood of bullshit is itself no more than a fractional part) is enormous ... (OT, 71-72).

As Frankfurt sees it, then, the "extraordinary prevalence and persistence of bullshit in our culture" (OT, 4) is a fact that cannot be so easily dismissed.

And yet, it is also an undeniable fact that our Western society is an advanced culture that is in many ways prospering – materially as well as culturally and intellectually. But if so, it follows from what we have said before that our society must be characterized by its members knowing collectively a great many truths. So apparently, it is possible after all for the

members of a society to share knowledge of many truths in spite of the fact that that very society is also being massively infected by various forms of deception. But how can this be? After all, we have just learned that widespread deception and bullshit constitutes a serious threat to our common quest for truth.

A schematic summary of this train of thought might be useful at this point:

- (1) A society must in order to prosper be founded on large quantities of truths that are generally known by its members.
- (2) Massive amounts of bullshit and deception, more generally, preclude large quantities of truths from being generally known.
- (3) Our society is seriously infected with bullshit and deception.
- (4) Hence, in our society at most only a few truths can be generally known (by (2) and (3))
- (5) Hence, our society does not prosper (by (1) and (4))
- (6) But our society does prosper.
- (7) Contradiction (by (5) and (6)).

Let us agree to call the problem to which this piece of reasoning gives rise “Frankfurt’s Puzzle”.

Frankfurt is not insensitive to the fact that he has reasoned himself into a corner:

We are all aware that our society perennially sustains enormous infusions – some deliberate, some merely incidental – of bullshit, lies, and other forms of misrepresentation and deceit. It is apparent, however, that this burden has somehow failed – at least, so far – to cripple our civilization (OT, 7).

Postmodernists and relativists in general could happily take Frankfurt’s puzzle as a *reductio* of the view that truth is at all important. Unsurprisingly, that reaction to the puzzle is passionately dismissed in *On Truth*:

Some people may perhaps take this complacently to show that truth is not so important after all, and that there is no particularly strong reason for us to care about it. In my view, that would be a deplorable mistake. (OT, 6-7)

Instead Frankfurt proposes to avoid the paradox by, in effect, restricting the scope of the second premise. Bullshit prohibits the truth from being known, unless, that is, people are generally good at detecting signs of fraud:

We can successfully find our way through an environment of falsehood and fraud, as long as we can reasonably count on our own ability to discriminate reliably between instances in which people are misrepresenting things to us and instances in which they are dealing with us straight. (OT, 72)

In other words, Frankfurt's answer to the question of how a thriving society is possible despite extensive bullshitting is that we – its members – are, or must be, generally good at detecting signs of deception.

This optimistic proposal, however, is apparently withdrawn one page later (OT, 73):

To be sure, we are rather easily fooled. Moreover, we know this to be the case. So it is not very easy for us to acquire and sustain a secure and justifiable trust in our ability to spot attempts at deception. For this reason, social intercourse would indeed be severely burdened by a widespread and wanton disrespect for truth.

If we are rather easily fooled and know this to be the case, we cannot after all to any great extent discriminate reliably between instances in which people are misrepresenting things to us instances in which they are dealing with us straight. In the end, therefore, Frankfurt provides no compelling explanation of how our society can prosper and at the same time be an arena for widespread deception and fraud, leaving us with no compelling solution to the puzzle that he has frame.

Is there any other solution to Frankfurt's puzzle? One assumption I will not question is that our Western society is indeed flourishing. It is certainly characterized by immense artistic and intellectual activity, modern science being among its most striking achievements. But what about the claim that a thriving society needs to be founded on truths that are generally known? To be sure, examples of prospering societies founded on bullshit and lies do not come readily to mind, but one may still question the claim that the truths have to be known by *all* or even *most* members for the society in question to prosper. Perhaps it is sufficient for that purpose that there is an enlightened minority of people firmly in tune with the facts of the matter, especially if they are in political power. In many Western countries, those occupying higher

offices are generally either knowledgeable themselves or, more importantly, have access to expert groups that are.<sup>9</sup> The fact that there exist simultaneously various subcultures – scientologists, astrologists, religious sects and so on – the members of which entertain beliefs that are in serious error does not threaten the prospects of material and intellectual wealth so long as those groups are small enough not to influence general elections.

The contention that our society is subject to extensive bullshitting could also be questioned, especially if it is taken to imply that the misinformation is evenly distributed. In reality, some areas of social life are more affected than others. The Internet is known not only for housing massive amounts of valuable data but also for providing a venue for dishonesty and sham. Science, by contrast, though by no means immune to fraud, contains mechanisms that at least seriously reduce its vulnerability in this regard.

Finally, there may be other reasons than those tentatively put forward by Frankfurt for dismissing the thought that widespread deceptive activity necessarily undermines the prospect of society converging on the truth. In the next section I will consider recent work in social epistemology suggesting that a society may thus converge even if only a fraction of its members are successful truth-seekers and truth-tellers, none of whom is capable of detecting signs of deception.

#### **4. A social epistemology perspective**

The problem of finding conditions under which the beliefs of the members of society converge on one and the same opinion has attracted attention in social epistemology. While a lot of work has been done in this area over the years,<sup>10</sup> the special case troubling Frankfurt – concerning the prospects of a communal convergence on the truth in an environment of extensive bullshitting – has, to my knowledge, been investigated only recently.

A directly relevant work is Hegselmann and Krause's 2006 paper "Truth and Cognitive Division of Labour: First Steps towards a Computer Aided Social Epistemology", which assesses the chances for the truth to be found and broadly accepted under conditions of what the author refer to as "cognitive division of labor" combined with a social process in which information is exchanged between the individual inquirers. By cognitive labor is meant that only some individuals are successful truth seekers as well as truth tellers. The rest are

---

<sup>9</sup> I believe this to be true of most of the old European democracies. The US under George W. Bush is a more complex case.

<sup>10</sup> The best known work in philosophy is probably Keith Lehrer and Carl G. Wagner's *Rational Consensus in Science and Society* from 1981. More recent work includes Galam and Mascovici (1991), Friedkin and Johnsen (1990), Nowak et al (1990), Duffuant et al (2000), Sznajd-Weron and Sznajd (2000), Stauffer (2004), and Lorenz (2005).



unreliable in the way they form beliefs, which makes them also, in Frankfurt's words, unfit for telling the truth. Being unreliable as informants, they are bullshitters in one sense of that word, even though they are not necessarily intentionally bullshitting. The social process consists in a mutual exchange of opinions between all individuals, truth-seekers or not.

H&K study this setting both mathematically and by means of computer simulations using an iterative procedure. Fortunately we need not go into much of the mathematical details. It suffices, for our philosophical purposes, to get a reasonable firm grasp of the main ideas and the results that follow. One assumption is that each individual starts out with a certain opinion that is taken to be arbitrarily chosen. For definiteness, we may think of it as a real number between 0 and 1 representing some quantity that is being assessed, e.g., the robustness factor in our previous engineering example. The opinions are then assumed to be made public for all to see upon which each individual updates his or her opinion based on the opinions of the others. At this point it is assumed, not unrealistically, that a given individual takes into account only the other opinions that are, from her point of view, are not too exotic. In other words, a given individual bases her new opinion on (i) her own old opinion, and (ii) those opinions that are within a certain distance  $\varepsilon$  from her own old opinion, where the number  $\varepsilon$  is referred to as the "confidence level". To be specific, the individual suspends judgment, as it were, between these views by taking on a new view corresponding to their average value.

How does truth enter into the picture? H&K assume that there is a true opinion,  $T$ , in the space of possible opinions that may be capable of "attracting" individuals in the sense that they have a tendency to approach it, perhaps because they are using rational argumentation, reasonable thinking, sound experimental procedures et cetera. Thus H&K – wisely, I believe – choose to abstract from the qualitative nature of the methods of inquiry, focusing instead on their quantitative reliability.

There is in this model an objective as well as a social component determining the opinion of a given individual. Objectivity is, as we just saw, captured by the degree to which a given individual is attracted to the truth. The social component amounts to specifying how much weight the individual assigns to the opinions of her "peers". It is helpful to take a quick look at the main equation to which these considerations give rise:

$$(HK) \quad x_i(t+1) = \alpha_i T + (1-\alpha_i) f_i(x(t)), \text{ with } 1 \leq i \leq n.$$

where  $x_i(t+1)$  is the new opinion of the  $i$ :th individual,  $\alpha_i$  the degree to which that individual is attracted to the truth and  $1-\alpha_i$  the degree to which her opinion is socially determined.

Setting  $\alpha_i > 0$  means that the  $i$ :th individual is to some extent attracted to the truth. Setting  $\alpha_i = 0$  means that there is no direct connect between her opinion and the truth but that her new opinion is rather the mere product of her own previous one and the opinions of her peers. Cognitive division of labor, in H&K's sense, results when only some individuals have  $\alpha_i > 0$ . H&K take care to point out that none of this should be taken to imply any conscious deliberative activity on the part of the individuals. In that sense, the equation assumes a reliabilist-externalist as opposed to an internalist perspective on inquiry.

Setting  $\alpha_i = 0$  turns the  $i$ :th individual into a kind of bullshitter, albeit one that is sensitive to social facts. Although the behavior of such an individual reflects a lack of concern with truth, she still adapts her own opinion to the views of her peers. An asocial bullshitter may be obtained as a special case by setting the confidence level equal to 0, meaning that she judges her own old belief to be the sole opinion worthy of serious consideration. Applying Peirce's Method of Tenacity,<sup>11</sup> an asocial bullshitter will simply stick to her own arbitrarily chosen initial position come what may.

H&K proceed to conduct their computer simulations with amusing and occasionally unexpected results. Let us start with the case where none of the individuals is a truth-seeker, and where the confidence level of each individual is fairly low so that only a few other opinions are taken into account. In other words, each individual is essentially applying the Method of Tenacity just alluded to. What will happen when the main equation is used repeatedly to update the opinions of the individuals is that a number of clusters are eventually formed consisting of individuals sharing the same opinion, each cluster being out of reach of, and hence incapable of influencing, the others. The convergence of one of those clusters on the truth will of course be a purely random affair. H&R describe the result of their simulations as "an eternal plurality of divergent views". (Fortunately, they choose not to draw any parallels to philosophical inquiry; that would have been discouraging.)

Suppose we assume, to take the other extreme, that all individuals are truth seekers in the sense of being attracted to the truth, if ever so slightly, and that circumstances are otherwise the same. Then what will happen is that there will be the same initial tendency for clusters to be formed, but these clusters will now, as the updating process continues, gradually approach the truth (and hence also each other), though they may not get to there in finite time.

These extreme scenarios still do not represent a case of cognitive division of labor. For such a case to arise some individuals have to be truth seekers and others not. The perhaps most important question, for our purposes, is whether or not many individuals have to be truth

---

<sup>11</sup> See Peirce (1877).

seekers in order for society as a whole to converge on the truth. Interestingly, the answer to this question is in the negative. H&K provide an example where 50 percent of the individuals are slightly attracted to the truth and where all others are social bullshitters. The result after some initial clustering is that the social exchange process finally leads to a consensus that is at least fairly close to the truth. This consensus includes the bullshitters who will, because of their social nature, become indirectly connected to the truth through the information they receive from reliable peers.<sup>12</sup>

Surprisingly, the position of the truth matters turns out to be of significance for this result. If the truth happens to be a more moderate position, this will increase the likelihood of communal convergence on the truth. If it happens to be an extreme position, the truth-seekers will still converge on it, and so will some of the bullshitters whose views were not too distant from the truth to begin with. Nevertheless, there will be an increased tendency for more distant bullshitters to form their own cluster far away from the truth. The result will be a split society where a minority of bullshitters having been left behind by the informed majority.

Another unexpected result is that as truth-seekers get better at their craft the chances of a communal convergence on the truth is reduced rather than – as one might have thought – increased. The effect is in fact easily explained. As the truth-seekers become more attracted to the truth they will tend to approach the truth more quickly. This also means that the likelihood decreases that they will influence a given social bullshitter. The effect will be a polarized society with a majority converging on the truth and one or more minorities approaching opinions distant from the truth. This will be so even if the truth happens to be a moderate view. The situation can be improved upon by making the bullshitters more social, i.e., by raising the confidence level of the individuals so that more opinions are seriously considered. Just as a fisherman can increase his chances of catching a fish by increasing the size of his net, so too a bullshitter can increase his chances of connecting to a truth-seeker by extending the range of views that he takes into account in forming his new opinion.

What drives many of these results is the fact that a bullshitter, while by definition lacking any direct connection to the truth, may become indirectly connected to it through a social exchange process. In our case, the social process dictates (i) that one should pay attention at least to views sufficiently similar to one's own and (ii) that one should "average" in the case of disagreement. These two social principles combine into the dictum to average in the face of peer disagreement.

---

<sup>12</sup> In their example, H&K have made the individuals slightly more social compared to their previous simulation by raising the confidence level just a bit, thus extending the range of other opinions that are taken into account when a given individual forms a new opinion.

As H&K are the first to admit, their model can be criticized from a number of perspectives as being simplistic and unrealistic. It assumed, for instance, that all individuals have access to the current opinions of all the others. What if the social exchange process has a network structure involving primarily local informational exchange? Moreover, averaging over the views of one's peers makes clear sense only in settings where the task is to assess the value of a quantitative variable. In many other contexts averaging has no clear meaning. Even when averaging does make sense, there are other ways of taking the views of one's peers into account, e.g. by taking the median value rather than the average value.

Another potential shortcoming of the model is that it does not accommodate the distinction between unreliable beliefs and unreliable reports. H&K assume that once an individual is reliable in her belief formation, she is also reliable in her reporting. The typical case of bullshitting would rather be one in which a reliable inquirer decides to report in an irresponsible and unreliable manner. Frankfurt's corruption of character hypothesis, which we encountered in section 3, makes sense only on the more complex model. Only on that model can the fact that you consistently decide to report unreliably gradually lower your reliability as a believer.

But I strongly suspect that the main philosophical point that can be extracted from H&K's work is still valid: There may be informational exchange process which, combined with the social pressure to average among peers, compensates for the fact that a part of the population lacks direct contact with reality by ensuring that that part nevertheless enjoys an indirect access to the way things are via the beliefs and reports of reliable peers, whose views they are forced to take into account. None of this assumes any special ability on the part of the individual inquirers to detect signs of fraud. *Pace* Frankfurt, communal convergence on the truth does not require that "we can reasonably count on our own ability to discriminate reliably between instances in which people are misrepresenting things to us and instances in which they are dealing with us straight" (OT, 72).

**Acknowledgements:** I wish to thank Mikael Janvid and Stefan Schubert for their comments on an earlier draft.

## References

- Armstrong, D. M. (1973), *Belief, Truth and Knowledge*, Cambridge University Press.
- Deffuant, G. Neau, D, Amblard, F, and Weisbuch, G. (2000), "Mixing Beliefs among Interacting Agents, *Advances in Complex Systems* 3: 87-98.

- Frankfurt, H. G. (2005), *On Bullshit*, Princeton University Press.
- Frankfurt, H. G. (2006), *On Truth*, Alfred A. Knopf: New York.
- Friedkin, N. E., and Johnsen, E. C., (1990), "Social Influence and Opinions", *Journal of Mathematical Sociology* 15: 193-206.
- Galam, S., and Moscovici, S. (1991), "Towards a Theory of Collective Phenomena: Consensus and Attitude Changes in Groups", *European Journal of Social Psychology* 21: 49-74.
- Goldman, A. I. (1976), "Discrimination and Perceptual Knowledge", *The Journal of Philosophy* 73: 771-91.
- Goldman, A. I. (1986), *Epistemology and Cognition*, Harvard University Press.
- Goldman, A. I., and Olsson, E. J. (to appear), "Reliabilism and the Value of Knowledge", in *Epistemic Value*, Pritchard, D. et al (eds.), Oxford University Press.
- Hegselmann, R., and Krause, U. (2006), "Truth and Cognitive Division of Labour: First Steps towards a Computer Aided Social Epistemology", *Journal of Artificial Societies and Social Simulation* 9 (3).
- Jones, W. E. (1997), "Why do we value knowledge?", *American Philosophical Quarterly* 34, No. 4, October: 423-439.
- Kvanvig, J. L. (2003), *The Value of Knowledge and the Pursuit of Understanding*, Cambridge University Press.
- Lehrer, K., and Wagner, C. G. (1981), *Rational Consensus in Science and Society*, Dordrecht: Reidel.
- Lorenz, J. (2005), "A Stabilization Theorem for Dynamics of Continuous Opinions", *Physica A* 335 (1): 219-223.
- Olsson, E. J. (2007), "Reliabilism, Stability, and the Value of Knowledge", *American Philosophical Quarterly* 44 (4): 343-355.
- Peirce, C. S. (1877), "The Fixation of Belief", reprinted in *Philosophical Writings of Peirce*, Buchler; J. (ed.), Dover: New York, 1955: 5-22.
- Pritchard, D. (2007), "Recent Work on Epistemic Value", *American Philosophical Quarterly* 44: 85-110.
- Riggs, W. D. (2002), "Reliability and the Value of Knowledge", *Philosophy and Phenomenological Research*: 79-96.
- Sosa, E. (2003), "The Place of Truth in Epistemology", in *Intellectual Virtue: Perspectives from Ethics and Epistemology*, Oxford University Press: 155-179.

- Sznajd-Weron, K, and Sznajd, J. (2000), "Opinion Evolution in Closed Community", *International Journal of Modern Physics C* 11: 1157-1166.
- Stafuffer, D. (2004), "Difficulty for Consensus in simultaneous Opinion Formation of Sznajd Model", *Journal of Mathematical Sociology* 28: 25-33.
- Swinburne, R. (1999), *Providence and the Problem of Evil*, Oxford University Press.
- Williamson, T. (2000), *Knowledge and Its Limits*, Oxford University Press.
- Zagzebski, L. (2003), "The Search for the Source of Epistemic Good", *Metaphilosophy*, 12-28.