

The Guru Effect

Dan Sperber

Published online: 6 March 2010
© Springer Science+Business Media B.V. 2010

Abstract Obscurity of expression is considered a flaw. Not so, however, in the speech or writing of intellectual gurus. All too often, what readers do is judge profound what they have failed to grasp. Here I try to explain this “guru effect” by looking at the psychology of trust and interpretation, at the role of authority and argumentation, and at the effects of these dispositions and processes when they operate at a population level where, I argue, a runaway phenomenon of overappreciation may take place.

Obscurity of expression is considered a flaw. Not so, however, in the speech or writing of intellectual gurus.¹ It is not just that insufficiently competent readers refrain, as they should, from passing judgment on what they don’t understand. All too often, what readers do is judge profound what they have failed to grasp. Obscurity inspires awe, a fact I have been only too aware of, living as I have been in the Paris of Sartre, Lacan, Derrida and other famously hard to interpret *maîtres à penser*. Here I try to explain this “guru effect.”

1 Believing and Trusting

There are two ways of holding beliefs in one’s mind. Holding a belief may be experienced—to the extent that it is experienced at all—as plain awareness of a fact, without awareness of reasons to take it to be a fact. So are held most of our ordinary beliefs. They are delivered by our spontaneous cognitive processes, the reliability of which we take for granted without examination. I believe that it is sunny because I see that it is; I believe that it rained yesterday because I remember that it did; and I believe that you are in a good mood because this is how I spontaneously interpret the

¹I am using here the English word “guru,” not the Sanskrit word from which it is derived.

D. Sperber (✉)
Institut Jean Nicod, EHESS-ENS-CNRS, 29, rue d’Ulm, 75005 Paris, France
e-mail: dan@sperber.com

expression on your face. Here, “because” introduces not reasons I might have weighted in forming these beliefs, but the causal processes through which I come to have them. Such beliefs are “intuitive” in the sense that they impose themselves on us without our being aware of the process through which they do so.

Other beliefs I hold because I also believe there is a good reason to hold them. I believe that it will be sunny tomorrow because so said the weather report, and I find its next-day predictions reliable enough. I believe that you just made up with your friend on the phone because this is the best explanation I can find for your suddenly improved mood. In these cases, “because” introduces a reason for my belief. Such beliefs are “reflective” in the sense that we entertain them together with the reasons we have to accept them.²

Entertaining a reason is as much a cognitive process as is perceiving, remembering or mood-sensing. Conversely, the fact that perception, memory and mood-sensing are reliable cognitive processes would give us a reason, if we cared for one, to accept the beliefs they generate. The contrast I want to draw between “reflective” and “intuitive beliefs” is not between beliefs held because of a cause and beliefs held because of a reason, but between beliefs held with or without mentally represented reasons.

Reasons to accept a belief may be “internal,” that is, have to do with the content of the belief: I believe some proposition because I accept an argument from which this proposition follows. Such an argument may be based on evidence: I believe that the cake in the oven is properly baked because the knife blade I inserted in it came out dry. The argument may be purely formal: I believe that there is no greatest prime number because, given any prime number however large, I know how a prime number greater than this one can be computed.

Reasons to accept a belief may also be “external,” that is, have to do with the source of the belief: I believe that what I have been told or what I read because I judge the source to be reliable. I believe my friend Mary will come to dinner tonight because she said she would and I trust her. I believe that there are tensions between the President and the Prime Minister because so says *Le Monde*, a newspaper I find reliable on such issues. Catholics believe that the Father, the Son, and the Holy Spirit are one because they trust the priests who tell them so.

The belief that a friend, the newspaper or the priest is trustworthy may itself be held intuitively or reflectively. I intuitively trust my friend Mary, without having ever reflected on her trustworthiness. When, on the other hand, a belief in a source’s trustworthiness is held in a reflective manner, it may, just as other reflective beliefs, be based on internal reasons having to do with the content of the belief or on external reasons having to do with the source of the belief. Christian children may believe the priest is trustworthy because their parents (whom they trust intuitively) told them he is—an external reason. I believe that *Le Monde* is, on the whole, trustworthy because I have had much direct evidence of this trustworthiness—an internal reason.

We may initially accept a person’s authority on the basis of her reputation—an external reason—, and then update our degree of trust on the basis of her record—an internal reason. I first went to doctor Z because she was warmly recommended to

² For the distinction between intuitive and reflective beliefs, see Sperber 1997.

me. Now I keep going—and I myself recommend her to others—because, in my experience, her diagnoses and advice have been confirmed, and have heightened my confidence in her.

2 Trusting and Interpreting

Updating the strength of one's trust in an authority figure may be affected by what, in the psychological literature on reasoning, is known as a "confirmation bias"³: reflecting on their beliefs and wanting to justify them cause believers to pay more attention to confirming than disconfirming evidence, thereby increasing their initial confidence. Internal evidence of trustworthiness is typically interpreted on the basis of prior trust. I followed doctor Z's prescription and got well in a week when I had hoped to be cured in three or four days. Still, I trust her and take the fact that I was cured as further evidence of her trustworthiness. If my trust in doctor Z had been wavering, I might have taken the fact that it took me a whole week to get well when it seemed reasonable to hope for a more rapid recovery as a reason to question doctor Z's trustworthiness.

The more the evidence is open to a variety of construals, the greater the risk of a confirmation bias. Few things better lend themselves to divergent construals than obscure statements. It is not surprising then often to find that their interpretation is strongly biased by the prior authority granted to their source. Divination practices around the world provide the best illustrations of this kind of interpretive charity: consultants interpret pithy statements—and the Pithia's own, in her time—in a way that is relevant to them and confirms the powers they attribute to the diviner:

Fortune-teller: I see a tall man...I see a bird...people you care about are in pain...Consultant: Amazing! Yes, everybody was sick after Thanksgiving, and the guy who sold me the turkey was very tall indeed.

Interpretive charity is not, however, an odd departure from normal interpretive practices, and not only pithy statements but all utterances leave room for interpretation. Quite generally, sentences vastly underdetermine their interpretation. Typically, they contain referring expressions the referent of which is not linguistically determined, they are multiply ambiguous, and they are open to a whole range of construals going from literal, to loose, to figurative. An utterance never fully encodes the speaker's meaning. Rather, it provides a richly structured piece of evidence from which the hearer (or the reader) can infer the speaker's (or the writer's) meaning. In this inferential process, hearers are helped by considerations of relevance. Utterances raise expectations of relevance that guide the comprehension process towards an interpretation that satisfies these expectations.⁴ So, for instance, if John arriving late tells me, "I missed the bus," I understand him to refer to the bus that could have brought him in time, and to mean "miss" in the sense of fail to arrive in time to board and not of feel sad about the absence of, of fail

³ Wason 1960. See Mercier and Sperber 2009 for a new approach to the confirmation bias.

⁴ This is a central claim of Relevance Theory (Sperber and Wilson 1995)

to hit with a projectile. In fact, typically, I home in on the contextually relevant interpretation without being aware of alternatives.

We expect what people tell us to be relevant, and we interpret it in a way that confirms this expectation. To the extent that speaker themselves expect us to home in on an optimally relevant interpretation of their utterances and produces utterances the optimally relevant interpretation of which is the very one they intended, what could be seen as an instance of the confirmation bias is, in this case, a rational way to achieve coordination and understanding.

Relevance itself has two aspects: everything else being equal, the greater the cognitive effects derived from the processing of an utterance (or, for that matter, any other type of information), the greater its relevance. For instance, if you want to know at what time is the next train to Manchester, it would be more relevant for you to be told “it is at 5:16” than to be told “it is sometimes after 5.” The more precise statement not only entails all the consequences of the vaguer one, but it also entails further consequences that you are likely to pay attention to: more cognitive effect, more relevance. The second aspect of relevance has to do not with cognitive effect but with processing effort. Everything else being equal, the greater the effort needed to process an utterance, the lesser its relevance. It would be more relevant for you to be told of the next train to Manchester, “it is at 5:16” than to be told, “it is twenty-two minutes after 4:54” (unless, of course, the lapse between 4:54 and the departure of the train is of special relevance to you) although the two statements are synonymous and carry exactly the same consequences. The second, more convoluted statement requires greater processing effort: more effort, less relevance.

So, we expect what we are told or what we read to be relevant, that is, to carry sufficient effect to be worth our attention and to do so without causing us unnecessary effort of comprehension. Of course, speakers or writers tend to overestimate the relevance of what they have to say, and hearers’ or readers’ expectations of relevance are frequently disappointed. In particular, when people of no particular authority express their thoughts in an obscure manner, we often revise down our already moderate expectations of relevance to a level where trying to make sense of what they say is not even worth the effort. On the other hand, when we trust that what we are told is relevant, the fact that some stretch of discourse or text requires more effort leads to the expectation it will carry more effect (extra effort being a price paid for extra effect, thus maintaining the overall level of relevance).

In fact, departing from plain and easy formulation is often a way of signalling that something other than plain meaning is intended. I cannot resist using a famous if somewhat exaggerated examples of Paul Grice. He writes:

Compare the remarks:

- (a) Miss X sang “Home Sweet Home.”
- (b) Miss X produced a series of sounds that corresponded closely with the score of “Home Sweet Home.”

Suppose that a reviewer has chosen to utter (b) rather than (a). (Gloss: Why has he selected that rigmarole in place of the concise and nearly synonymous sang? Presumably to indicate some striking difference between Miss X’s performance and those to which the word singing is usually applied. The most obvious supposition is that Miss X’s performance suffered from some hideous defect.

The reviewer knows that this supposition is what is likely to spring to mind, so that that is what he is implicating).

Grice 1989: 37

This example illustrates how a deliberately opaque formulation directs one towards a richer interpretation.

In other cases, comprehending an utterance may involve extra effort but in a way that was not intended, or at least not overtly intended. It is as if the speaker or writer had no easier way to express herself or as if she expected greater ease of comprehension on the part of her hearers that they are actually capable of. Even so, if the speaker or writer chose to go ahead and express a thought hard for her audience to understand, she is thereby suggesting that the thought in question is relevant enough to be worth the effort.

As children we were often told things that we didn't quite understand but were clearly intended to. Little Lucy is told by her teacher that cucumbers are 95% water (an example I borrow from Andrew Woodfield). She thinks of water as a liquid. Now, cucumbers are solid, not liquid objects; water does not flow out of them; so what could the teacher mean? Accepting, however, the authority of the teacher, Lucy now believes, without fully understanding it, that, somehow, cucumbers are 95% water. The very difficulty of grasping this idea indicates to her that this is a relevant piece of information, worth remembering and thinking about until she can make better sense of it.

Lucy was also told by her parents and at Sunday school that God is everywhere. This too she believes with only partial comprehension. Whereas many children end up understanding how solid bodies such as cucumbers can mostly be made of water, the belief that God is everywhere remains impossible to fully comprehend. This mysteriousness is, if anything, even better recognized by theologians than by children. Given that, for the faithful, the relevance of the belief is beyond question, its very mysteriousness is a strong indication of its significance. Impenetrability indicates profundity.

In front of religious mysteries (divine omnipresence, the Trinity), believers stand in awe. They may derive some relatively unproblematic consequences from these beliefs (e.g., divine omnipresence implies that there is no place to hide from God) but it takes theologians to aim at sophisticated interpretations that, anyhow, are never final. For most believers, the existence of mysteries is, in fact, more relevant than their actual content. Because of the authority they grant religion, believers are convinced that the content of mysteries would be extraordinarily relevant to them if only they could grasp it. The fragmentary interpretations of mysteries that lay and clerical believers arrive at are wholly guided by this certainty of relevance. The existence of barely glimpsed hyper-relevant content is yet another confirmation of the supreme authority of religion.

The writing of many philosophers, especially but not uniquely in the so-called continental tradition, is full of hard-to-understand passages where difficulty is presented as pertaining not to expression but to content itself, as being not a rhetorical device but a direct and unavoidable aspect of sophisticated thinking. Here are a few characteristic quotes (which, being cited out of context, are not here to be judged, let alone sneered at; still, no contextualisation would make them simple and easy to understand):

- “Beauty is a fateful gift of the essence of truth, and here truth means the disclosure of what keeps itself concealed.” Martin Heidegger

- “Consciousness is a being, the nature of which is to be conscious of the nothingness of its being.” Jean Paul Sartre
- “In a culture like ours, long accustomed to splitting and dividing all things as a means of control, it is sometimes a bit of a shock to be reminded that, in operational and practical fact, the medium is the message.” Marshall McLuhan
- “If *différance* is (and I also cross out the ‘is’) what makes possible the presentation of the being-present, it is never presented as such. It is never offered to the present. Or to anyone. Reserving itself, not exposing itself, in regular fashion it exceeds the order of truth at a certain precise point, but without dissimulating itself as something, as a mysterious being, in the occult of a nonknowledge or in a hole with indeterminable borders (for example, in a topology of castration).” Jacques Derrida

The point I am trying to illustrate is independent of the quality and clarity of what the authors had in mind when writing these passages. Maybe, each and every one of them had in mind an important thought that could not be expressed in any simpler way. Maybe some readers (including, possibly readers of the present essay) have grasped these thoughts and been illuminated by them. The fact is that, for most if not all readers, the interpretation of such statements is highly problematic. Still, the very effort required tends to be seen as an indication of high relevance and to favour interpretations consistent with this indication. If they cannot come to any clear and plausible interpretation, readers may nevertheless seek fragmentary and tentative interpretive hypotheses that go in the expected direction. Even if these statements remain hopelessly opaque, readers may take their very opacity as evidence of their depth.

Faced with an inordinately recondite statement, readers have the choice between a negative judgment: the author had no good reason to be obscure, and a positive explanation: the author wanted to convey a thought too deep for plain and simple expression. With a prior high confidence in the intellectual worth of the author, negative judgment is almost ruled out and depth can be inferred, even if no satisfactory interpretation of the statement in question is ever arrived at. Prior appreciation of an author justifies a positive construal of difficult passage. So far, so good. Things may go wrong if, in a viciously circular manner, this construal is taken as further justification for the appreciation.

3 Authority and Argumentation

Suppose there is a claim that you won’t accept just on my authority, I may still try to convince you by providing a reasoned argument, starting from premises that you are willing to accept (because you already believe them or because, for them, my authority is sufficient), going through a series of steps the validity of each you can judge by yourself, and concluding with the claim I want you to accept. The logical force of an argument does not depend on the authority of whoever puts it forward. A mathematical proof expounded by a known swindler may be convincing all the same. While there is no sure way to tell by mere inspection a true statement from a false one (unless the false statement is self-contradictory or contradicts what is

already known to be true), competent examination is enough to tell a valid from an invalid argument. Thus, when authority fails to provide a sufficient external reason for accepting a claim, argumentation may provide an appropriate internal reason.

Authority and argumentation seem to be two quite different paths to persuasion, and, to a large extent, they are. From an evolutionary point of view, the capacity to produce and evaluate arguments might have emerged as a way of partially overcoming the risk of deception and manipulation involved in accepting the authority of communicators.⁵ Historically, the transition to modernity can be described as the replacement of authority by argument as the main basis of justified beliefs. In intellectual style, there is often a clear opposition between those who trust more authority than argument, and those who trust more argument than authority. Nevertheless, in communicative practices, what we find is not a dichotomy between appeal to authority and appeal to reason, but a variety of interactions and overlaps between the two forms.

To begin with, trivially, authority can be argued for. For instance, in John, 14: 11, Jesus says: “Believe me when I say that I am in the Father and the Father is in me; or at least believe on the evidence of the miracles themselves.” Jesus gives a reason to accept his authority to those who would not spontaneously do so.

More interestingly, trust in authority may give us a reason to accept the validity of an argument without examining its steps, or even without quite comprehending it. Bobby fails to understand the math teacher’s demonstration that there is no greatest prime number, but the very fact that the teacher presents what she claims is a proof causes Bobby to accept as a proven fact that there is no greatest prime number—and he is right, of course! This extends to non-demonstrative arguments. For example, people looking for an effective weight loss program may stumble on the following argument: “Where is the scientific evidence that eating the controlled carb way is healthy? By adhering to a controlled carbohydrate nutritional approach, an individual who chooses to eat nutrient-dense foods ... is more likely to meet his nutritional needs, promoting good health, than he would by following a calorie-restricted, fat-deficient diet. ... For studies that support the health benefits of a controlled carbohydrate nutritional approach, click here. All these studies confirm that not only is controlling carbohydrate consumption effective, it actually results in improved health parameters” (<http://atkins.com>). Even if they do not fully comprehend the argument or are not able to weigh its force and cannot be bothered to click and look at the additional evidence proposed, people may be swayed by the fact that what looks like a forceful argument is being put forward for their consideration.

Suppose that you accept on trust some argument of mine as valid and its premises as true. Then, of course, you also will accept as true the conclusion of the argument. Given however that you are just relying on my authority, should the fact that I have argued for this conclusion give you a better reason to accept it than if I have merely asserted it? After all, if you are willing to take my word for the soundness of an argument, why not just take it directly for the truth of its conclusion? Well, the very fact that I produced an argument, even if you are unable to assess its validity, or cannot be bothered, is of relevance to the evaluation of its conclusion. To argue is to

⁵ See Sperber 2001

make an effort in order to appeal to one's audience's reason. It can be seen as a mark of respect for one's audience (just as to refuse to argue is a mark of disrespect). A valid argument is harder to fake than a true statement. To argue is to expose oneself to critical examination. So, the very fact that I made the effort and took the risk involved in arguing may contribute to the believability of my conclusion, even if the argument remains unexamined.

When paying with a check, you may offer to present some identification: sometimes, this very offer will be seen as evidence of your trustworthiness and will be declined just because it was made, while, if not spontaneously offered, identification would have been requested. Of course, swindlers know this too and can use apparent forthrightness to achieve devious goals. Similarly, the apparently honest display of argumentation can be used to impress, browbeat, or even deceive one's audience, and had been developed as a rhetorical technique by the Sophists depicted in Platonic dialogues.

My interest here, however, is in honest rather than dishonest gurus. Honest gurus are not trying to deceive their audience. Nevertheless, they may produce arguments that will persuade most of their readers not by their logical force, but by their very difficulty. A recent illustration is provided by *The Emperor's New Mind* by the eminent physicist Roger Penrose (1989). As summarized by the blurb of the book, Penrose "argues that there are facets of human thinking, of human imagination, that can never be emulated by a machine. Exploring a dazzling array of topics—complex numbers, black holes, entropy, quasicrystals, the structure of the brain, and the physical processes of consciousness—Penrose demonstrates that laws even more wondrously complex than those of quantum mechanics are essential for the operation of a mind" (my italics). Given the wealth of premises from different fields of knowledge and the complexity of the argument, I doubt that most readers are in a position to evaluate what, if anything, Penrose demonstrates. Still, coming from such an authoritative source, the very elaborateness of the argument is enough to suggest that it can withstand a level of scrutiny that most readers would be quite unable to provide, and that Penrose is offering a hard-to-grasp but plausible and highly relevant perspective on the relationship between fundamental physics and human psychology.⁶

4 Runaway Trust in Authority

A possible explanation of the obscurity of a statement made by an authoritative source may be that it expresses some important thought that could not be formulated in a simpler way. Similarly, a possible explanation of the difficulty of an argument may be that there is no simpler way to justify its conclusion. When my only alternative is to question the otherwise well-established authority of the source just because I have trouble understanding it, these explanations may be the best I can come up with, and, if so, I should accept them. Such "inferences to the best explanation" may in turn justify my accepting a statement as true or an argument as valid, even though I don't quite understand them. On the other hand, how could my

⁶ For doubts that it is so, see Dennett 1989.

failure to fully understand a statement or an argument ever justify me in granting even more authority to its source? Obscurity need not be held against an author—after all, it may just reflect the limits of my own understanding—, but how could it be held in her favour? An obvious risk in upgrading the authority of a source because of the obscurity of its pronouncements is that of running into the vicious circle I mentioned: the favourable interpretation I give of an obscure text is based on the prior authority I grant its source; if I then use this interpretation to value up this authority, and then this enhanced authority to interpret even more favourably the next obscure text from the same source, a string of obscure texts (or, for that matter, reinterpretations of just one of them) might cause me to grant near-absolute authority to a source just because I don't understand it. Are individuals on their own predisposed to commit this kind of fallacy? I see no reason to believe they are, or at least, not systematically. On the other hand something of the sort happens in the collective recognition of authorities.

Authority is social relationship that involves at least two individuals, and typically many more. Authority in a group goes with reputation. The reputation of a person is the more or less consensual view of her competence and reliability that spreads through repeated acts of communication across a social group. Individuals may just state that So-and-so is knowledgeable or wise, or they may give examples of this knowledge and wisdom. They may also discuss the interpretation and the value of specific pronouncements. Clear statements and easy arguments may become the objects of a collective evaluation, but only obscure statements and difficult arguments are likely to become the objects of a collective endeavour of interpretation.

As long as the interpretation of a text is not settled, its evaluation is likely to be based on external rather than internal criteria. We don't know what X meant in making some obscure statement, but, given the authority we recognise him, we have reasons to think that he was expressing a very important idea. In fact, if we did not think this, we would not be involved in trying to comprehend that statement. Participating in a collective process of interpretation amounts to publicly vouching for the value of what is being interpreted. Moreover, it seems sensible to take the amount of attention paid to thinkers and their thought as a rough indication of their importance—and it would be sensible if it were not for the fact that these individual evaluations build on one another and together spiral into ever greater devotion.

Participants in a collective process of interpretation have a double stake in the value of the text they are working on and in the authority of its author. The greater this value and this authority, the more they are justified in joining the process, and the less the tentative and partial character of their interpretations can be seen as a negative reflection on their own intellectual abilities. Moreover, participating in such a collective process involves not just an intellectual but also— and more surely— a social benefit, that of belonging, of getting recognition as a person in the know, capable of appreciating the importance of a difficult great thinker. Not participating, on the other hand, may involve the cost of being marginalised and of appearing intellectually stale and flat.

Here emerges a collective dynamics typical of intellectual schools and sects, where the obscurity of respected masters is not just a sign of the depth of their thinking, but a proof of their genius. Left on their own, admiring readers interprets

one recondite passage after another in a way that may slowly reinforces their admiration (or else render them wary). Now sharing their interpretations and impressions with other admirers, readers find in the admiration, in the trust that other have for the master, reasons to consider their own interpretations as failing to do justice to the genius of the interpreted text. In turn these readers become disciples and proselytes. Where we had the slow back-and-forth of solitary reading between favourable interpretation and increased confidence in authority, now we have a competition among disciples for an interpretation that best displays the genius of the master, an interpretation that, for this purpose, may be just as obscure as the thought it is meant to interpret. Thus a thinker is made into a guru and her best disciples in gurus-apprentices.

Unlike the people in Andersen's tale pretending to admire the emperor's non-existent clothes, participants in the collective dynamics of guruification need not be, and generally are not in bad faith: they have strong external reasons for their appreciation—reasons that they provide one another—, which in turn lead them to favourable interpretations that provide them with further internal reasons. Moreover they need not even be wrong: human intellectual history is full of challenging propositions and arguments that turned out to be true and important. Still the epidemiological mechanism⁷ I have briefly sketched explains how many obscure texts and their authors come to be overestimated, often ridiculously so, not in spite but because of their very obscurity.

References

- Dennett, Daniel (1989). Murmurs in the cathedral (review of R. Penrose, *The Emperor's New Mind*). *The times literary supplement*, September 29–October 5, pp. 55–57.
- Mercier, H., and Sperber, D. 2009. Intuitive and reflective inferences. In Evans, J. St. B. T. and Frankish, K. (Ed.) *In two minds: Dual processes and beyond*. Oxford University Press.
- Penrose, R. 1989. *The emperor's new mind: Concerning computers, minds, and the laws of physics*. Oxford: Oxford University Press.
- Sperber, D. 1996. *Explaining culture: A naturalistic approach*. Oxford: Blackwell.
- Sperber, D. 1997. Intuitive and reflective beliefs. In *Mind and Language* 12(1): 67–83.
- Sperber, D. 2001. An evolutionary perspective on testimony and argumentation. *Philosophical Topics* 29: 401–413.
- Sperber, D., and D. Wilson. 1995. *Relevance: Communication and cognition*, 2nd ed. Oxford: Blackwell.
- Wason, P.C. 1960. On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology* 12: 129–140.

⁷ See Sperber 1996