

Believing the implausible: rational error, cognitive bias and spontaneous social theory

Declan Gaffney

'Perhaps most striking in these data is the fact that, on all measures, the vocabularies of the children from advantaged homes were larger than the vocabularies of the parents in the welfare families.' US Department for Health and Human Services 1999

<http://mentalhealth.samhsa.gov/schoolviolence/part1chp9.asp>

'Hart and Risley also found that the vocabulary of a child of three from the professional class was greater than that of the adult from the family receiving benefits.' Chris Ruane MP Hansard 21 Jan 2004 : Column 499WH

'By age three upper-income toddlers not only had vocabularies twice as large as the welfare children: they also had bigger vocabularies than the welfare parents.' Dick Mendel, 'The most effective child development programmes work with kids and their parents' in *American Prospect* 2 November 2004

'Even more remarkably, in the same work by Hart and Risley (1995) three-year old children from professional families used a wider vocabulary than parents receiving welfare.' David Miller, 'What is social justice?' in Pearce N and Paxton W (ed.) *Social justice: building a fairer Britain* (IPPR London 2005)

'A three year old from a professional family usually has a larger vocabulary than the average adult in a welfare household.' [from a post on the Freerepublic.com site dated 29 August 2007]

“A three year old from a professional family usually has a larger vocabulary than the average adult in a welfare household.” I don't think so. This sounds like elitist bullshit to me.' [from a reply to the above post]

1 Believing the implausible

People often believe things that, given the evidence available to them and their own cognitive abilities they really shouldn't believe. They sometimes

unthinkingly assent to, or fail to query, statements which a moment's consideration would show to be utterly implausible. It's not just that people get things wrong, but they get them wrong when they could reasonably be expected to get them right, and the patterns of error are often non-random; that is, people often don't just get things wrong in any old way but systematically get them wrong in specific ways.

Assent to false statements requires an additional level of explanation to that required for assent to true statements (the truth value of a true statement can standardly enter into the explanation for people believing it; the truth value of a false statement can't). Theories of how these sorts of errors arise are interesting in themselves; they can also play an important corrective role, placing us on our guard against the ways in which our own intellectual weaknesses, lack of attention and biases can lead us into error. And because widespread error can have causal impacts on the real world, they may play a role in explaining social and historical phenomena.

2 Rational error

One aspect of explaining error turns on how human psychological capacities may depart from what would be required for perfect rationality. For example, humans are widely believed to be congenitally bad at understanding basic probability (compared to say arithmetic or geometry) and this implies that their assessments of risks and their understanding of statistics are subject to severe error. Let's call this the limited capacity line of explanation.

Another line of explanation starts from the costs involved in arriving at reasonable judgements. We don't submit every statement we come across to rational scrutiny because that would be impossible, indeed irrational. So we assent to, or fail to query, implausible statements because their implausibility isn't glaring enough- to *us*- to attract our attention and encourage us to put the effort in to query them. We are more likely to notice and query blatant implausibility the more we are concerned with the subject, that is the higher the cost *to us* of accepting a wrong statement as true. Let's call the latter the economic line of explanation.

If the costs of being wrong are low enough or the costs of more intense scrutiny are high enough, assenting to or failing to query even a patently dumb proposition may be a case of *rational* error. Note that the limited capacity and economic lines both feed into the notion of rational error. If we're congenitally bad at judging probabilities then it requires more effort on our part to judge statements that turn on probability, so there would need to

be proportionally higher costs attached to *not* subjecting these statements to scrutiny before we would be inclined to put in the work. Contrariwise, if we were all good spontaneous probability theorists, the costs of scepticism would be much lower, tightening the margins of concern.

3 Cognitive bias

What I have referred to as rational error can help explain why people acting and thinking rationally don't apply fully rational criteria to judgments all the time. But there are also other, potentially non-rational factors at work. Studies of cognitive bias have identified a number of patterns of bad judgment: self-serving beliefs, just-world theories, fundamental attribution errors and so on. The unacknowledged motivations behind these patterns are easy to hypothesise, when they aren't already written into the definition of the bias in question. Thus self-serving beliefs protect self-esteem ('I got an A grade because I am clever and work hard; I got an F grade because I was unlucky in the questions on the paper that day'). Actor-observer bias and fundamental attribution error involve an explanatory asymmetry where the actions of *others* are explained in terms of ascribed dispositions and choices, ignoring the situations under which the actions were taken ('people like that just don't mind living like that'), while the role of the situation is stressed in accounting for one's own behaviour ('do you think I would live like this if I had any choice?'). The bias known as 'just-world theory' ('the world can't be *that* unfair') may mitigate the stress of contemplating painful situations and the fear of finding oneself or loved ones in similar situations by explaining them in terms of the behaviour, capacities or preferences of those affected (as in the brutal example of the blame-the-victim bias of juries in rape trials).

Cognitive bias and rational error can pull in opposite explanatory directions: some cognitive biases may be more likely the *more* involved we are in a situation, while rational error is (by definition) more likely the *less* involved we are. This alerts us to a paradoxical feature of judgment. The ideal of the 'impartial observer' as used by Adam Smith for example suggests that those who have the least personal investment in the truth, falsehood or probability of a statement are those who are most likely to give a correct judgment on its truth-value. But impartiality in the sense of the absence of a personal stake will not be enough to guarantee good judgment in all cases. We often believe the dumbest statements about things which are of no significance to us at all, just *because* they are of no significance to us. As we are neither perfect nor omnipotent reasoners, even our judgments on matters in which we are impartial are subject to rational error.

4 Socially situated error

So we have two lines of explanation for why people often assent to, or fail to dissent from, dumb propositions: rational error and cognitive bias. Note that neither of these lines of explanation appeals to notions such as ideology, false consciousness, deliberate misinformation or any kind of social brainwashing. Phenomena of that type may (or may not) have a big influence on the things that people believe or go along with. But they aren't our concern here, where we want to look at how people come to assent to dumb statements through what we can think of as the normal, universal processes of judging factual statements which humans are generally equipped for. (Some would say that the appeal to universality here is itself ideological.)

Nonetheless rational error and cognitive bias share with theories of ideology the fact that they associate incorrect judgment with social situations. For example, the self-esteem which an individual seeks to protect through self-serving beliefs or actor–observer bias is inherently socially situated, so the effects of bias on particular judgments can be expected to vary- at a statistical level of analysis, I stress- between individuals in different situations and also to show some similarities between individuals in similar situations. Likewise, whether it is rational to assent to an implausible statement depends on the relative costs of assent and of scepticism to the individual concerned, and as noted above, these will vary with, amongst other things, the importance of the content of the statement to that (socially situated) individual. Thus rational error is no less socially situated than cognitive bias. We may all have propensities to both, but their effects – including the type of statements that we will subject to stringent scrutiny or let go – are likely to vary in non-random ways with individuals' social situation.

Another potential area of overlap with theories of ideology is that socially situated error may have *causal* effects. This can arise for example if a majority of individuals shared certain biases or thresholds of concern with regard to certain types of statement: the very widespread acceptance of various forms of pseudo-scientific racism in the late 19th and early 20th century is a privileged example.

The key difference with theories of ideology lies in the account of the origins of error. Theories of ideology tend to hold that false beliefs are functional for ruling classes or élite groups and that this functionality somehow explains their hold on those who are *not* members of the ruling class or élites. The function of ideology is to prevent subordinates recognising their real interests. A theory of socially situated error makes no functionalist assumptions, nor

does it introduce the philosophically dubious notion of 'real' interests. Nonetheless, while it is not functionalist, it does help explain how defective judgment can be 'functional' in the strictly negative sense that when the costs of holding to false beliefs rise for individuals, those beliefs will become increasingly *dysfunctional* for them and should therefore become more open to challenge.

5 The inescapability of theory

Consideration of ideology brings us to a third line of explanation for error: the inescapability of theory. Beliefs entail beliefs: nothing that we believe stands in complete isolation from other beliefs we may hold, and many things are implicitly linked to whole sets of other beliefs. Everyone is a spontaneous theorist trying to maintain coherent relationships between their various beliefs, or at least trying to avoid incoherence becoming a problem.

The rational way to deal with incoherence, for example when a new fact disturbs existing patterns of belief, is to revise beliefs, or so we like to think: but revising beliefs is costly, and may therefore be *irrational*. Ways of avoiding large scale belief revision include dismissal of awkward facts; compartmentalisation, that is failing to follow through the entailments of statements; and forgetfulness, as when people persistently return to errors which they have previously been led to recognise as errors. For current purposes, spontaneous theorisation can help explain some cases where people assent to dumb statements: this is more likely when the statements in question, dumb as they are, don't entail major revision of other beliefs *or* when they can be seen to be implied by other beliefs.

We can see that the idea here is somewhat broader than that of ideology, but it shares with that concept a stress on the interconnectedness of beliefs. We tend to think of ideologies as structured and relatively coherent, if mistaken, sets of beliefs. The idea of socially situated error we are advancing here is far less committed to structure and coherence but does ascribe a role to relations of entailment between beliefs in generating and, no less importantly, permitting error.

Ideology is sometimes held to be collective rather than an individual phenomenon. Like many others, I have difficulty making sense of this idea. Socially situated error has a collective aspect insofar as bad judgment on some matters is likely to vary systematically according to people's social situation. Moreover, error is likely to be self-reinforcing once it is widespread enough within a reasonably coherent social group, as we are less likely to challenge

error when it is widely shared among people we know (the costs will often be higher, amongst other things). Thus we can get from the individual to the collective without invoking a notion of the collective as anything other than the result of individual, socially situated error.

Moving from the individual to the collective in this way means that interactions between individuals come to play an explanatory role. When error is widespread, the explanation can be expected to involve not just the individual factors which explain error in each individual case but the ways in which each individual judgment takes account of what the individual believes about the beliefs of those with whom he is socially involved. (Economists will recognise that this raises issues for formal modelling.)

If we put these three lines of explanation for error togetherⁱ, we can see that all we are appealing to is some very basic, even banal psychology- analogous to the 'folk psychology' which economists use. Humans have certain cognitive capacities, but they aren't perfect; they prefer to expend less effort on things which are of less importance to them; they have interests which affect their judgment; they are theorists in the sense that one of their (imperfect) cognitive capacities is the ability to generalise and follow logical entailments; and they take account of what they believe others believe in forming judgments.

6 Spontaneous social theory

Social and economic inequality are fertile areas for our dispositions to biased judgment to flourish, as any background theories we hold about how inequality arises are likely to have implications for ourselves. For example, self-serving motivations may encourage us to believe that whatever our earnings, they are certainly no *more* than the appropriate rate for our own combination of talent and effort. This bias in favour of the justice of the rewards we receive can still function even if people think that the distribution of rewards in our society is very unjust. A middle class professional may well deplore the fact that he is paid five times as much as a domestic cleaner, while continuing to believe that *his* salary results from his talent and effort. (See how he reacts if you agree with him on the injustice of the cleaner's situation and go on to add 'and of course it's even more unjust that you're paid so much when you aren't even that smart or hard-working'.) While we may not believe that rewards are always or even usually scaled to contribution, if we want to believe that our own rewards are the results of our own efforts and talents, we are committed to some kind of background theory about the relationship between rewards, talents and efforts, however vague that theory

might be: the vaguer the better, perhaps. Note that while such a theory has normative implications, it is not itself a normative theory.

It doesn't take much reflection to see that generalising from our own far from perfect judgments about the relationship between our *own* capacities and our *own* economic rewards to a general theory of social inequality is unlikely to result in a good theory. This doesn't really matter to us, as most of us are not reflective theorists most of the time. We are spontaneous social theorists, not out of choice but because most beliefs tend to have at least some logical entailments which even at our most unreflective it is difficult to evade. Unless we are already inclined to (or paid to) reflect on such matters, even blatantly implausible statements which fit in well with spontaneous social theories influenced by unreflective biases may well get under our radar without challenge.

That may not be particularly important to us as individuals: why should we care if we have a few not very deeply-held beliefs which are a bit dumb, as long as they aren't about things that are important in our lives? Aren't we just in the same position as the character in Donald Barthelme's story who went through life thinking that the mark of Zorro was an 'N'? But as we indicated above, it may be important *socially* if a lot of other people have similar propensities to similar sorts of error about similar subjects, because that can create an environment in which certain types of dumb statement become widely accepted, and widespread acceptance of dumb statements can have causal effects.

7 The infant mother

A few years ago it was widely reported in the United States that 'three-year old children from professional families used a wider vocabulary than parents receiving welfare.' This claim appeared in a couple of speeches by members of the Bush administration and travelled far and wide, and still crops up today, as can be seen from the quotations at the start of this note.

To anyone who has ever met a three year old from a professional family, this would be a truly alarming statement. At least it would be if the statement got under their cognitive radar without attracting suspicion. Alternatively, people might decide to do some basic checking. Five minutes of Googling would lead them to a note by the linguist Geoff Nunberg showing that the study cited, by Hart and Risley, said nothing remotely similar to what had been claimed.

Maybe you've guessed it already: what that study showed (amongst other important things) was that toddlers from professional families used a wider vocabulary in talking to their mothers than mothers from welfare families used *in talking to their toddlers*. The study didn't look at the everyday working vocabulary of the mothers: it was solely concerned with the mother-child interaction. The mothers in the study, whether professional or on welfare, didn't attempt to use their full vocabularies in talking to their toddlers. Presumably they recognised, like most reasonable adults, that the laudable principle of not talking down to kids has certain practical limits.

Geoff Nunberg concluded 'I can't hear this sort of thing without being reminded that the word infant comes from a Latin word that means "not having the power of speech." In that sense, the claim that welfare mothers have thousand-word vocabularies is infantilizing in the literal sense of the term. It suggests that the mothers don't actually have enough language to be able to make sense of the world they live in. And from there it follows that these women simply aren't in a position to articulate their needs or make reasoned judgments about their lives -- that they literally can't speak for themselves.'

The claimed equivalence in vocabulary between mothers on welfare and middle class three-year olds was nonsense, and pretty obvious nonsense at that, but that didn't stop it getting into circulation on both sides of the Atlantic. The question is why people were prepared to believe and repeat the claim that a large section of the adult population of the U.S. had the linguistic competence of toddlers, while for others the claim was so obviously ridiculous (see the last of the quotations at the head of this paper). And that brings us back to the propensity for error in thinking about social inequality.

8 Some hypotheses

As a line of explanation of the fact that this idiotic proposition got under the cognitive radar of people who ought to know better, I would hypothesise that many people are motivated *not* to put a lot of effort into querying statements that (1) support the view that differences in social and economic status arise primarily from differences in effort and in capacity, *as long as* those statements are (2) concerned with people who they don't know and (3) who are situated somewhere *below* their own economic and social level.

The basis for Condition (1) has already been explained (section 6) , and Condition (2) is obvious: the costs of assenting to false statements concerning (types of) people you know are likely to be much higher than if they concern

the sort of people you are unlikely to know. Condition (3) arises from the perception that statements about those *above* one's own level engage different types of motivation than statements about those below one's level. Consider the statement 'Chief executives are paid so much more than managers because that's *exactly* how much smarter and harder working than managers they are'. Self-serving motivations may well incline managers to query statements of *this* type, asking about the role of luck, or irrelevant capacities (for sycophancy, say) or privilege. But there is no need to invoke cognitive bias to explain why managers would balk at this statement: it's already a dumb statement, and all that is necessary is that people shouldn't be motivated *against* holding it up to scrutiny, as chief executives might be, for example.

Now nobody, probably, really believes that capacity and effort explains everything about social and economic inequality any more than anybody really believes they explain nothing. But when people are living in dire poverty- a situation which they are unlikely to have chosen, after all- statements that imply that their poverty is due to a basic lack of capacity may well fit in well enough with the spontaneous social theories of enough other people to gain an unthinking acceptance. And as the poverty of mothers on welfare in the United States is pretty extreme, statements positing an extreme lack of capacity, even going well beyond what is plausible to the point of ascribing infantile capacities to these mothers, may gain some credence even among those who morally deplore poverty.

We can easily see how this type of message might be ideologically useful to a particular administration, but the point I would like to stress is that where that kind of explicit motivation isn't present- and it certainly wasn't among all of those who retailed the myth of the infant mother- thinking about the sources of error can help us understand what's going on when smart people believe dumb statements. Rational error (weak grasp of probability combined with the costs of being wrong being low), cognitive bias (just world theory and attribution error), and spontaneous social theory (the myth didn't challenge but rather supported wider beliefs about how society works) all played their role in this case, or so I suggest.

6. Conclusion

Of course you may disagree. You may feel that the 'infant mother' statement isn't *so* implausible that we need to theorise about how people came to go along with it. And you'd have a point, or at least half a point, because plausibility is itself an inherently subjective concept, and if a lot of people don't see a statement as implausible, then it sounds reasonable to argue that it just *isn't* implausible. To which my reply is that while plausibility may be an inherently subjective concept, it is also inherently normative: it makes sense to say that people are mistaken in taking statements as plausible when they aren't, even when a great many people are doing so. It makes sense to say that they *should* have known better, that they should have been prepared to put more work in, that they should have been more impartial, been prepared to revise other beliefs if necessary and so on. Of course there is a judgment (mine) involved in saying that the statement is implausible, but if I'm wrong, it's because my judgment is defective, not because it's in some logical or semantic sense wrong to regard widely accepted statements as implausible.

I chose this example not just because it struck me as a spectacular instance of assent to the implausible but also because of the importance I ascribe to underlying spontaneous social theory in explaining that assent. The question raised for me by the 'infant mother' myth, and many other contemporary myths about poverty, is this: what kind of spontaneous social theory, however crude, is implied by assent to this kind of dumb statement? For the fact that the statement was not immediately seen as implausible by people who could be expected to know better is a *social* fact. It tells us something, I would argue, about how large numbers of people see the world we live in and the place of inequality in that world.

In this sense, some types of error are social phenomena, arising from the convergence of different individuals with propensities to error on particular false statements. And we should take this seriously, because when false statements are able to flourish, they can have causal impacts on the world in virtue of which they are, in fact, false.

ⁱ I want to put to one side situations where people somehow *make* themselves believe implausible factual statements because this is in their conscious interest: for example, when politicians bring themselves to believe implausible claims about the military capabilities of foreign countries because they need to believe this in order to be able to claim to be acting in good faith in promoting a war. This kind of self-deception raises dizzying psychological and philosophical problems which have been explored by the late Donald Davidson among others, and this is not the place to dive into those deep waters. Nonetheless there is bound to be a thin line between rational error and self-deception, in the sense that it may be hard to distinguish between cases where people assent to implausible statements because they are better off going along with them than querying them, and cases where they in some way deliberately force their own judgment because they *believe* they will better off that way.