The Fragmentation of Reason: Précis of Two Chapters[1]

STEPHEN P. STICH
Rutgers University

My four distinguished critics have made my task a bit easier by aiming almost all of their comments at two of the chapters in *The Fragmentation of Reason*. In those chapters I offer arguments against two widely accepted accounts of what it is for one cognitive strategy to be better than another. The first account gives pride of place to the analysis of our ordinary notions of cognitive evaluation—notions like *justification*. The second links cognitive assessment to truth. Though there is a fair amount going on in these two chapters, their central arguments share a common structure. I'll begin by sketching that structure and saying something about how I would like the arguments to be understood. This will, I hope, be of some help to readers who are not familiar with the book. It will also provide a useful backdrop against which I can set out my responses to some of my critics' objections.

My preoccupation with the evaluation of cognitive strategies began when I learned about the work of psychologists like Nisbett, Ross, Tversky, Kahneman and others whose experiments seemed to show that bright people in unthreatening environments often do a very bad job in a variety of reasoning tasks. A very natural reaction to this work is to wonder how we might improve matters. Are there steps we can take that will enable people to do a better job of forming and revising their beliefs? This question suggests two rather different lines of inquiry. One is largely empirical: What sorts of interventions will succeed in changing the way people go about the business of reasoning? The other is normative: What sorts of changes would be *desirable*? What is it that makes one strategy of

---

reasoning *better* than another? These normative questions are center stage in the second half of *The Fragmentation of Reason*, and it is my rejection of the two most commonly endorsed answers that my fellow symposiasts find most disturbing.

Before setting out those answers and saying why I reject them, let me propose a thought experiment that I have often found useful in explaining the view I want to defend. When it comes to actually designing ways of improving reasoning, all sorts of practical constraints are, of course, going to be important; but the goal of the thought experiment is to allow us to ignore these for a while. Imagine that we are confronted with an enormously powerful Genie who offers to make one of several changes in the way our minds work. Each of these changes will alter the way we form and update our beliefs—they will change the way we reason. Our job is to decide which of these changes we would find desirable. To help us with our decision, the Genie tells us something about the consequences of various options. One proposed change will result in our acquiring more justified beliefs and fewer unjustified ones than we would if we were to retain our current strategies of reasoning. Another change will result in our acquiring more true beliefs and fewer false ones than we would as things stand now. Still other alternatives will be described later on. One way of thinking about the desirability of various strategies of reasoning is to ask which of the Genie's offers we would be inclined to accept. If, for example, we really do value "reliable" (or truth generating) cognitive strategies, then the change which will leave us with a more reliable cognitive system should be genuinely tempting.

It's my contention that, when we view the matter clearly, most of us are not going to find the offer of a cognitive system that generates more true beliefs to be especially tempting, nor are we going to find any special attraction in cognitive systems that yield more justified beliefs and fewer unjustified ones. At first blush these are rather startling claims. But the arguments that support them are, I think, both straightforward and persuasive. Those arguments begin with the familiar distinction between intrinsic and instrumental value (or desirability). Presumably we will accept the Genie's offer to transform our cognitive system in a way that will produce more justified beliefs (or more true ones) only if we think that having justified beliefs (or true ones) is either intrinsically desirable, or instrumentally desirable, or both. Yet, I claim, for most of us they are neither. In defending this claim, the central step is to show that both justification and true belief are highly *idiosyncratic* notions—each is just one member of a large family of more or less related notions that happens

to have been picked out by our own cultural tradition, despite having no obvious advantages over many of the other notions in its family.

Alvin Goldman's analysis of the notion of justification that is embedded in everyday thought and language[2] (*our* everyday thought and language) provides a convenient way of illustrating this idiosyncracy. According to Goldman, justified beliefs are those produced by psychological processes that "would result in a truth ratio [i.e., a ratio of true to false beliefs] that meets some specified high threshold."[3] But, as Goldman notes, this account requires that we specify something about the nature of the world in which the processes are operating. And, on his view, the account of justification that "best accords with our intuitions"[4] is one that assesses the reliability of psychological processes not in the actual world but in what Goldman calls "normal worlds," which he defines as those in which our general metaphysical and ontological beliefs are true.[5] For current purposes, it is not important whether Goldman is right. What is important is to see that if he is even close to being right, then *our* concept of justification occupies a small and rather idiosyncratic region in a large space of more or less similar concepts that can be generated by altering the specification of possible worlds in which the reliability of psychological processes is to be assessed. A much larger space of justification-like concepts can be generated by varying other parameters in Goldman's account. But there is no reason to think there is anything at all special about the region of this space in which our intuitively endorsed concept of justification happens to be located. The fact that we have inherited *this* idiosyncratic concept of justification rather than one of the others is simply a cultural and historical accident.

Now it is my belief that when people see clearly just how accidental and idiosyncratic our notion of justification is, most will recognize that there is nothing intrinsically desirable about having justified beliefs or cognitive processes that tend to produce them. Of course, it may still be instrumentally desirable to have justified beliefs, since such beliefs may do a particularly good job at leading to something else that people find valuable. And, given the wide variety of things that people value, there could hardly be a general argument showing that justified beliefs cannot possibly be instrumentally valuable. Still, the idiosyncrasy of the notion of justification—the fact that it is just one member of an enormous family

---

[2]   *Epistemology and Cognition* (Cambridge, Massachusetts: Harvard University Press, 1986), pp. 58-59.
[3]   Ibid., p. 106.
[4]   Ibid., p. 107.
[5]   Ibid.

of justification-like notion—makes it *prima facie* implausible that having justified beliefs will be more instrumentally valuable than having beliefs which accord with any of the other justification-like notions. Those who claim that having justified beliefs is instrumentally optimal for attaining some goal owe us an argument for their claim. And none of the arguments I know of along those lines are at all persuasive.

Imagine, now, that the Genie offers us a choice between two different ways of modifying our cognitive systems, and tells us that the first option will result in a higher ratio of justified to unjustified beliefs while the second will result in a higher ratio of JUSTIFIED* to UNJUSTIFIED* beliefs, where JUSTIFICATION* is some other notion in the justification family. Since having justified beliefs is no more intrinsically valuable than having JUSTIFIED* beliefs, and since we have no reason to think that having justified beliefs is more instrumentally valuable than have JUSTIFIED* beliefs, I would not know how to choose.

The argument aimed at showing that having true beliefs is neither intrinsically nor instrumentally desirable has much the same structure. The essential step is to show that our notion of true belief is highly idiosyncratic. The easiest route to this conclusion begins with the assumption that belief tokens are brain state tokens. However, unlike some other brain state tokens, beliefs have semantic properties. This is because there is an intuitively natural way of mapping beliefs to propositions (or truth conditions, or possible states of affairs). A belief token is true if it is mapped to a true proposition (or a possible state of affairs that actually obtains, etc.) The idiosyncrasy of our notion of true belief results from the idiosyncrasy of our intuitive strategy for mapping beliefs to propositions. Part of that strategy exploits the existence of elaborate causal chains linking the concepts out of which our beliefs are built to various objects, kinds and classes in the world. Consider, for example, the belief token that I would express by saying "Thales drank water". Our intuitive "interpretation function" maps one component of that belief to a certain ancient sage, and another component to H_2O. Thus the belief is true if and only if that sage drank H_2O. However, there is another function, albeit a counter-intuitive one, that maps the first component of the belief to some other ancient, and still another function that maps the last component to H_2O or XYZ. If we exploit this second, counter-intuitive, function then we can define what might be called the TRUTH* condition for the belief token I express by saying "Thales drank water." That belief token is TRUE* if and only if a certain ancient sage drank either H_2O or XYZ.

From this point onward, the argument against the desirability of true belief is completely parallel to the argument against the desirability of

justified belief. Our intuitively sanctioned interpretation function, like our intuitively sanctioned notion of justification, occupies a small region in a large space of alternatives. And in light of the rather complex and quirky causal links that our intuitive interpretation function seems to demand, it is hard to think that there is anything special about this particular region in interpretation-function-space. The fact that we have inherited this idiosyncratic interpretation function rather than some other one is largely a matter of cultural and historical accident. Thus, for most of us, there will be nothing intrinsically desirable about having beliefs that are mapped on to true propositions (or to the actual world) by the idiosyncratic intuitive interpretation function. Moreover, as Lycan notes in his contribution to this symposium, there are bound to be alternatives to the intuitive notion of true belief that are instrumentally superior—if we had a few more TRUE*···* beliefs and a few less true ones we would surely be better off.

Though this is only a bare bones sketch of my arguments against using the intuitive notions of justification and true belief in evaluating cognitive strategies, I hope it succeeds in conveying at least a rough idea of how those arguments work.