
TRUTH, JUSTIFICATION, AND THE AMERICAN WAY

BY

KENT BACH

He is no Lex Luthor but Steve Stich is up to serious mischief in *The Fragmentation of Reason*.¹ He means it when he says that “truth has lots of competition” (FR, p. 117) and that justification does too. So if someone doesn’t bring a halt to the wave of dastardly deeds perpetrated in this book, epistemological anarchy could prevail. Various philosophical crimefighters² have tried to nail the culprits operating within the book’s organized mob of arguments, but they have netted only the small fry, not the big fish.

Stich conducts a dual assault, on truth and on justification. He argues that having true/justified beliefs is neither intrinsically nor instrumentally desirable. “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 cultural tradition, despite having no obvious advantages over many of the other notions in its family” (SS, pp. 180–181). Indeed, “there are bound to be alternatives to the intuitive notion of true [or justified] belief that are instrumentally superior” (SS, p. 183). It may be a jungle out there, but truth and justification are the Mom and apple pie of epistemology. They must be protected.

1. *The Assault on Truth*

It is important to realize that Stich is not making any radical claims about truth itself. Although he labels his position “pluralist pragmatist relativism,” he is not contending, as this name might suggest, that there are many kinds of truth, that truth is a matter of what works, or that one

person’s truth is another one’s falsehood. He is claiming, rather, that there is nothing special about truth as a cognitive value, hence no special reason to assess cognitive processes, methods, and practices for how well they conduce to true beliefs. Mind you, Stich is not endorsing a policy of ignorance and error or suggesting that truth is no better than falsehood. His contention is more subtle than that. He contends that truth has many rivals, TRUTH*, TRUTH**, etc., each determined by a different “interpretation function,” and that truth is not inherently superior to its rivals. For the interpretation function for truth is on a theoretical par with the others; when assessed on pragmatic grounds, truth is sure to fare better than some members of the TRUTH* squad but, equally surely, worse than others. In short, truth is nothing special.

In defense of truth I will focus on four critical assumptions Stich makes, two explicitly and two implicitly. He explicitly assumes (1) that aiming at having true beliefs involves valuing truth and (2) that we have intuitions about interpretation functions (including the “intuitively sanctioned” one). I will argue, contrary to (1), that valuing true beliefs is not really a matter of valuing truth, and that (2) mistakenly presupposes that we have access to our “belief-brain-states” other than in terms of their truth-conditional contents. Stich’s implicit assumptions are (3) that the interpretation function for TRUTH* (e.g.) can be specified independently of the one for truth, and (4) that there are effective strategies for pursuing alternatives to truth. Against (3) and (4) I will argue that he has not indicated how to “get at” TRUTH*, either conceptually or practically. He has provided no reason for supposing that TRUTH* and company are viable and not merely notional alternatives to truth. Exposing these four assumptions will show that Stich’s denunciation of truth rests on a levels confusion about the role of truth in our thinking.

BELIEF AND VALUING TRUTH

When Stich says “valuing true beliefs is a profoundly conservative thing to do” (FR, p. 118), a natural rejoinder is this: “I should stop valuing true beliefs only if what you say is true. But if what you say *is* true, I should *not* let it affect what I value or believe.” However, this *ad hominem* argument exposes only the awkwardness in advocating Stich’s position, not a problem inherent in the position itself. It does not get to the heart of the issue, which concerns what we do when we do what he describes as “valuing true beliefs.” He seems to assume that in forming beliefs we aim at truth, as if we adopt some and not others because they possess a certain virtue. As a result, his arguments that truth is neither intrinsically nor instrumentally valuable (FR, pp. 118–124) seem misdirected. The provocative title question of Chapter 5 of FR, “Do We Really Care

Whether Our Beliefs Are True?", which Stich answers with a resounding "No!", is a red herring. He seems to suppose that we aim at true beliefs because and only because we value having true ones, as if truth were merely incidental to belief. This suggests that we could form beliefs independently of aiming at true ones. However, we don't decide to form a belief because it possesses some special feature, such as being pleasant, being agreeable, being provocative—or being true. Forming beliefs is not like choosing apples from a produce bin.

Gilbert Harman has provided a more plausible and perspicuous way of characterizing the connection between what we do in forming a belief and what we aim at in so doing: we want to believe something only if it is true. For instance, "I want my beliefs about people's telephone numbers to be true" (SS, p. 196). Harman uses the word "true" here but interestingly, though he takes no note of this, in his illustration "true" drops out: I want to "believe that Steve's telephone number is 683-1396 only if Steve's telephone number is 683-1396." Nor does it occur in his schematization: "[I often] desire that I believe that p only if p ." Indeed, this schema is needed if Harman's point is to be made without using the word "true." Aiming at the truth means wanting to believe something only if it is true, but this cannot be redescribed as "wanting to believe something only if something." Perhaps Stich's idea that aiming at true beliefs involves valuing truth is just an artifact of this curious linguistic fact. At any rate, thanks to the schema, "wanting one's beliefs to be true" can be rephrased in a way that makes no use of "true" and no mention of truth.³

A schema similar to Harman's is deployed by Paul Horwich, in his recent book *Truth*,⁴ to show that truth doesn't really enter into what one wants. He develops a "minimalist" account of truth, so called because not only does he endorse the "equivalence thesis," that the belief/proposition that p is true iff p , he rejects the need for anything deeper or less trivial about the "nature" of truth. Contending that the point of having the concept of truth is to apply it to items whose content one does not or in some cases cannot produce, he defends a deflationary conception of truth and argues against the need for a weightier notion, correspondence (or "robust") truth.⁵ Even if Horwich's minimalism is mistaken—Harman recognizes a need for both an "immanent"⁶ and a "transcendent" notion of truth—Stich, since he supposes that interpretation functions must be invoked to explain what truth is, just takes for granted that truth is inherently and primarily a transcendent or correspondence notion. As Harman argues, Stich's critique is aimed at the wrong target. Considering that a belief that p is true (in the deflationary sense) iff p , Harman finds that "The notion of truth here is simple. . . . Stich takes truth to be a much more complicated matter. But the complexity he discusses is not so much a complexity concerning truth as

it is concerning when it is possible to attribute to a person the belief that p " (SS, p. 196).

I would go further than Harman. The point is not merely that the notion of truth here is simple but that it doesn't enter into what we do when we form beliefs. When we consider what to believe with respect to p and proceed to believe that p (rather than believe not- p or just withhold belief), we do not judge the belief that p to be true or even the proposition that p to be true. We just judge that p , pure and simple. Forming a belief does not involve semantic ascent.

The complexity Harman alludes to arises in connection with translation or interpretation, which is a matter (in his terminology) not of immanent but transcendent truth. Stich confuses the task of giving a third-person account of the contents of someone else's language or belief-brain-states (BBSs) with possessing a language or a "language of thought." From the third-person point of view relevant to translation or interpretation, one is concerned with correlations between expressions or BBSs (syntactically or neurally specified) and reference, satisfaction, or truth conditions. The enterprise of translation or interpretation presupposes access in *nonsemantic* terms to the items being translated or interpreted. One need not engage in this enterprise in respect to one's own language in order to think one's own thoughts and speak one's own mind.

To appreciate this point, think of the transcendent/immanent distinction not as a distinction in kinds of truth (after all, how many kinds are there?) but as a distinction in levels, corresponding to the distinction between metalinguistic and object levels.⁷ If I wonder whether "Snow is white," "Der Schnee ist weiss," or "#*\$ @ §%¶" (a mock sentence of Mentalese) is true, I mean true in English, true in German, or true in Mentalese, as the case may be.⁸ However, if I wonder whether my statement/belief that snow is white is true, I am really concerned with the propositional content (presumably nonlinguistic) of the statement/belief. I am speaking *in* English or thinking *in* Mentalese (let's suppose), and asking or wondering whether snow is white.

So it is misleading and unnecessary to describe valuing true beliefs as wanting to believe something only if it is true. Thanks to the equivalence schema, this can be put in a way that avoids any mention of truth: (for any substituent for ' p ') wanting to believe that p only if p . What we are primarily interested in is not tracking the truth but tracking the ways of the world. Stich's focus on valuing true beliefs seems to stem from a levels confusion. As speakers/thinkers we are not in the business of trying to correlate our words or ideas with items in the world. Everyday speaking and thinking, unlike interpretation, is not a matter of making such a correlation but of implicitly exploiting the one that obtains in one's own case.

INTUITIONS AND INTERPRETATION FUNCTIONS

Even though Stich casts his main question in pragmatic terms, “Do we really care whether our beliefs are true?”, he has an underlying theoretical concern. When he contends that “truth has lots of competition,” he means that there is nothing theoretically special about truth, not merely that there is nothing distinctively worthwhile in pursuing it. Supposing that truth is determined by a certain interpretation function that “maps belief-brain-states to propositions,” he proceeds to argue that if there is one such function “there are indefinitely many. Obviously, not just any mapping will do.” The right one, according to Stich, is the one that “capture[s] our intuitive judgments about the content or truth conditions of the mental states in its domain” (FR, p. 23). This prompts him to ask, “what is so special about those intuitions?” He contends that the interpretation function determining truth is not importantly different from the ones that determine TRUTH*, TRUTH**, etc. Truth is nothing special because the interpretation function for truth is “limited⁹ and idiosyncratic,” on a theoretical par with all the different mappings for TRUTH* and company.

The claim that the interpretation function for truth is but one among many, each determining a different rival to truth, should not be confused with indeterminacy of interpretation. Stich is not claiming that belief contents are inherently indeterminate or that there is no fact of the matter about them. Although he introduces the notion of an interpretation function as a pairing of psychological states (specifically belief states) with propositions or truth conditions (FR, p. 105), he is not claiming that there is a multiplicity of pairings which constitutes distinct but equally good theories of content. Indeed, he is willing to concede that beliefs do have determinate truth conditions, that there is a unique interpretation function which specifies them, and that “some completely unproblematic story can be told about propositions and what makes them true” (FR, p. 104). So he is not claiming that there are equally good but incompatible truth-conditional theories of mental content but that there are equally good TRUTH*-conditional (or TRUTH*...*-conditional) theories of content. In other words, the claim is not that different interpretation functions are equally good as accounts of truth but as accounts of properties that are equally good, namely truth conditions, TRUTH* conditions, . . . , and TRUTH*...* conditions.

According to Stich, the interpretation function for truth is distinguished merely by the fact that it is the “intuitively sanctioned” one. It is the one that articulates “our intuitive strategy for mapping beliefs to propositions” (SS, p. 182). He is content to argue rhetorically against the authority of our intuitions, the ones that sanction the mapping for truth. It may be tempting to defend their authority, but to me the real problem

is his undefended assumption that we even have intuitions about the truth conditions of our belief-brain-states (BBSs). By his own admission, “For my arguments to work, it is going to be important that the interpretation function invoked in an account of truth conditions for beliefs is constrained in a fairly serious way by commonsense intuition” (p. 106). But do we really have intuitions about truth conditions for our BBSs in the way we do for sentences or utterances?

With linguistic items we can associate a certain shape or sound with a certain proposition and think of one that its truth condition is provided by the other. With BBSs, however, we have no access to their neural “shape” and are thus in no position to *pair* them with propositions; we have no recourse but to identify them by their propositional contents. So it comes as no surprise that when Stich tells his story about interpretation functions, he starts with an account of how sentences in a natural language get their truth conditions and then, playing along with the language of thought hypothesis, supposes that “we can view belief tokens as neurally encoded inscriptions of the relevant well-formed formulas,” so that the question at hand becomes “how we can assign truth conditions to these cerebral inscriptions” (FR, p. 109). Whatever the answer to this question—many rudimentary theories have been bandied about lately—it is certainly not a question on which we have any commonsense intuitions. We simply have no access to these cerebral inscriptions.¹⁰ We cannot identify our BBSs in syntactic (assuming we have a language of thought) or in neural terms but only by way of their contents. We can individuate the BBS that is the belief that *p* as the belief that *p*, but surely what Stich means by the “intuitively sanctioned” interpretation function isn’t the identity function. Other than that, he has provided no reason for supposing that intuition is equipped to sanction *any* interpretation function. So whereas he finds it “hard to see why anyone would take our intuitively sanctioned interpretation function to be special or important” (FR, p. 23), I find it hard to see why anyone would suppose that people have *any* intuitions about the truth conditions of their mental states, except in the trivial way just mentioned.

THE TROUBLE WITH TRUTH*

Stich maintains that truth and TRUTH* are each determined by, indeed defined by, interpretation functions. Now if this were true of truth, truth *would* be nothing special—our understanding of what truth is would depend on knowing which interpretation function defines it (not that we have intuitions about this). However, truth does have some distinctive properties. For one thing, it possesses the disquotational property. For another, we know what *truth* functions are and how to grasp truth-functional connectives. If “and” is construed TRUTH*-functionally,

would “ p is TRUE* and q is TRUE*” entail “(p and q) is TRUE*” and vice versa? Similarly, there are straightforward elementary facts about the (truth-) functional roles of beliefs in combination, e.g., that believing two propositions leads to believing their conjunction and that believing a conditional and believing its antecedent leads to believing its consequent.¹¹ What analogous functional roles, if any, belong to beliefs characterized TRUTH*-conditionally? At any rate, with knowledge of such distinctive features of truth as the ones mentioned above, we don’t need to have intuitions about its interpretation function to know what truth is.

It is hard to see how TRUTH* could have analogous features (not that Stich suggests they do), so the question arises, “What is TRUTH*?” We know that it is different from truth but how does it differ from TRUTH**, TRUTH***, etc.? The only answer available to Stich seems to be that TRUTH* is determined by a distinct interpretation function. But which one? He shouldn’t rest content with specifying the different rivals to truth schematically, by means of different numbers of asterisks. Even if he insisted that what he calls TRUTH* is just a dummy, a stand-in for any arbitrary rival to truth, each determined by a different interpretation function, he would still need to face the question of how interpretation functions are to be specified. And here it wouldn’t do to specify a particular interpretation function by enumeration, by a list of BBS-proposition pairs. The problem with enumeration is that the list for TRUTH* would be infinitely long, and a partial specification of that list would not distinguish it from the (infinitely) many alternatives to TRUTH* that share at least those pairs. So if we want to know what TRUTH* is, we need to know how, at least in principle, to generate recursively the entire list for TRUTH*. The interpretation functions for TRUTH* and its cohorts must each be finitely specifiable so as to be distinguished from one another and not be dismissed as mere notional alternatives to truth.

Stich grants that there is nothing problematic about giving a truth-conditional theory of content, say by using the device of T-sentences to define truth recursively, à la Tarski, by means of a “theory that entails an infinity of theorems [T-sentences] of the form ‘ S is true if and only if p ’” (p. 107).¹² (Here “ S ” is replaced by a specification of a sentence, be it a sentence in a natural language or one in the head, a syntactically specified BBS.) Similarly, TRUTH* (or TRUTH**, etc.) could be defined by a theory whose theorems (T*-sentences) take the form “ S is TRUE* iff q ” (here “ q ” rather than “ p ” is used to indicate that the TRUTH*-conditional theory may—it need not in every instance—pair a given sentence with a different proposition than does the truth-conditional theory). This poses a problem for Stich. Notice that both truth conditions and TRUTH* conditions are specified in terms of propositions: S is true iff p and S is TRUE* iff q . But propositions determine truth conditions,

and if TRUTH* conditions are specified in terms of items that determine truth conditions, it would seem that there is something special about truth—unless, of course, there is an independent way of specifying TRUTH* conditions. So Stich needs to explain how TRUTH* conditions can be specified other than by enumeration or in terms of items with truth conditions.¹³ He needs to show that they are determined in a way analogous to the way in which truth conditions are determined, namely, compositionally.

Stich appreciates this when he suggests the interpretation function for TRUTH* be grounded in an “alternative set of base clauses” (FR, p. 116). The idea is to give REFERENCE* or SATISFACTION* conditions for the primitive singular and general terms in a language, be it Portuguese or Mentalese. Stich compares reference to its rivals REFERENCE*, REFERENCE**, etc., which he describes as “alternative schemes¹⁴ for nailing words on to the world” (FR, p. 115). He illustrates how, in the case of names and natural kind terms, one or two such alternatives might be sanctioned by intuitions. Reference, which is sanctioned by *our* commonsense intuitions, is accounted for by a causal theory.¹⁵ REFERENCE*, perhaps sanctioned by the commonsense intuitions of some alien group, might be accounted for by a cluster-theory. So whereas the name “Jonah” refers to Jonah, it may REFER* to someone (or something?) else. Stich imagines that there is no end to such alternatives to reference, each accounted for by a different theory and sanctioned by a different set of intuitions. Unfortunately, giving them different numbers of asterisks does not show them to be viable alternatives.

Stich himself illustrates how hard it is to avoid routing REFERENCE* through reference and TRUTH* through truth. He says, of “the belief token I would express by saying ‘Thales drank water,’ . . . there is another [interpretation] function, albeit a counterintuitive one, that maps the first component of the belief to some other ancient, and still another function that maps the last component to H₂O or XYZ” (SS, p. 182). But how has Stich specified what “Thales” and “water” are mapped onto? He hasn’t really specified the “other ancient,” and if he did, he would have to use an ordinary referring expression to specify him, i.e., to refer, not REFER*, to him. He has specified the REFERENT of “water,” but only by using “H₂O” to refer to H₂O and “XYZ” to refer to XYZ (we all know what that is). Indeed, to state that the “belief token is TRUE* iff a certain ancient sage drank either H₂O or XYZ” is to give its TRUTH* condition in terms of a (non-standard) *truth* condition. This example illustrates a recurring theme: however compelling Stich’s position may be in the abstract, it loses its force once one tries to make it concrete (an outcome symptomatic of an occupational disease in philosophy, known to pathologists as *schematosis*). Stich has not shown

that alternative semantic properties like TRUTH* and REFERENCE* can be specified independently of truth and reference.

PURSUING TRUTH

This theoretical problem has practical consequences. A bona fide improvement on truth must be capable of being characterized systematically if it is to be pursued. The problem is evident, for example, in Lycan's characterization of "gruth," which is determined by the interpretation function that maps one's BBSs onto "whatever set of propositions-believed would have been optimal given the believer's circumstances and life history" (SS, p. 202). Unless there is a way of characterizing the function that determines gruth systematically and in advance of achieving the gruth, one could never be in a position to recognize the virtue of pursuing it. As we will see, merely arguing that improvements on truth exist in the abstract does not show that there is some identifiable improvement that we can pursue.

Stich thinks people could "lead happier, healthier, more rewarding lives" than they do with "cognitive systems [that] tend to produce true beliefs" (FR, p. 127). Truth may be better than some (maybe most) of its rivals, but there are some that are better than it. Perhaps there is one, TRUTH!, that is the best of all. There is provably a more modest improvement on truth, which may be called GRUTH (to borrow Lycan's term). Let q be some true proposition, say q , which one believes but whose negation one would be better off believing. Let GRUTH be the same as truth except that not- q is GRUE, so that all of one's true beliefs except the belief that q are GRUE and one would be better off holding the GRUE belief that not- q . Unfortunately, one would have to know which proposition this is in order to pursue GRUTH rather than truth. A further trouble with GRUTH is that it is defined in terms of a *specific* belief.¹⁶ What Stich is dangling before us is something more ambitious, the possibility of at least one alternative to truth, call it TRUTH!, whose *systematic* pursuit would make us better off than pursuing truth. What might that involve?

If TRUTH! is something whose systematic pursuit would make us better off than we are by pursuing truth, we need to know how to identify instances of it. Stich, despite his pragmatic protestations, is conspicuously silent when it comes to identifying strategies for pursuing TRUTH!. His Chapter 6, "A Pragmatic Account of Cognitive Evaluation," concerns the assessment of cognitive systems (human and others) and urges a consequentialist approach, but it offers nothing on how to overcome the shortcomings in one's cognitive performance attributable to aiming at truth. Simply advising inquirers to aim at TRUTH! would be as vacuous as advising investors to buy low and sell high. In order

to offer more, Stich would have to confront the problems mentioned previously, of showing how we could have access to our BBSs other than in terms of their truth-conditional contents and how TRUTH! conditions are to be specified independently of truth conditions. For example, if the TRUTH! condition of a certain BBS is q , it wouldn't do to identify this BBS as the belief that p or as the one whose *truth* condition is p . Unless it is shown how we could have a direct line on TRUTH!, that is, access to the TRUTH! conditions of our beliefs independently of their truth conditions, there will be no effective strategy for pursuing TRUTH! and nothing practical about Stich's professed pragmatism.

Indeed, even if TRUTH! could be suitably specified, there would remain certain practical obstacles to pursuing it. The first two are variations on "ought" implies "can." (1) Considering the extent to which much of our cognitive system may be modular and thereby impenetrable, it might not be feasible to adopt cognitive strategies that would require fundamental change in these less than malleable subsystems. (2) Lycan argues from an evolutionary standpoint that "there is an obvious sense in which, though adaptive, [TRUTH!] is not desirable: [TRUTH!] is not a viable design option. . . . [Because of] the vagaries and vicissitudes of people's life histories, [Mother Nature] could not have done much better than to satisfice by making our belief-forming methods truth-conducive instead" (SS, pp. 204–205). Finally, (3) Stich really has not shown that the shortcomings of our belief systems are specifically attributable to valuing truth (as opposed to any of its rivals). What limits the fulfillment of our goals and leads to failure and frustration is not that our beliefs aim at being true but, thanks to our propensity for error, that so many of them are false. And ignorance, the absence of true belief, is arguably even more deleterious to our health (we often risk error to overcome ignorance). I don't know how to measure their respective adverse consequences, but Stich has not begun to show that valuing true belief is in the same league with error and ignorance.

So there are serious obstacles to reaping the benefits of some new and improved truth value. Practical considerations aside, the real problem, if Alvin Goldman is correct, is that Stich has simply confused the issue: "Stich tends to conflate the question of what I.F. we accept with the question of what truth value we prefer. . . . Surely, acceptance of the ordinary I.F. carries no commitment to a preference for truth over error" (SS, p. 191), much less TRUTH!. I would put it this way: theories of content are one thing, cognitive strategies another.

2. *The Assault on Justified Belief*

Stich construes both reflective equilibrium accounts of justified inference

and reliabilist theories of justified belief as analyses or explications of our concept of justifiedness. Without defending his reading of Nelson Goodman to be proposing reflective equilibrium not (just) as evidence for but as constitutive of justified inference, he proceeds to devote an entire chapter (FR, Ch. 4) to this reading. But surely any analysis of justified inference that relies on reflective equilibrium is a non-starter. For if reflective equilibrium were constitutive, then what is a person supposed to be assessing when he assesses an inference as justified? That he likes it? That it sounds good? Stich quotes Goodman as proposing, “A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend” (FR, p. 77), but accepted or rejected in respect to what? If taken as constitutive, the reflective equilibrium test is too schematic to discriminate between good sound inference and good sounding inference.¹⁷

Stich debunks the efforts of analytic epistemologists on the grounds that the concepts they examine are idiosyncratic and of no particular value. Citing Goldman’s concession that the commonsense conception of epistemic justification is not only vague but multifarious, he suggests that to study these parochial conceptions is to engage in no more than an exercise in “domestic cognitive anthropology” (SS, p. 209). However, in my view Goldman should not have inferred non-uniqueness from vagueness. The vagueness in our commonsense notion of epistemic justification indicates merely that we have no *precise* concept of it. As Goldman’s own discussion in Chapter 5 of *Epistemology and Cognition*¹⁸ makes clear, justifiedness is a matter of degree and different standards and criteria of justifiedness reflect that fact. But it does not follow that these diverse ways of refining justifiedness or making it more precise show there to be more than one commonsense concept of justifiedness in the first place. Besides, there are other words for other notions, like “rational” and “well-grounded.”¹⁹ So Stich’s conclusion that it is “capricious or perverse to have an intrinsic preference for justified beliefs” rests on the unsupported premise that “our notion of justification is just one member of a large and varied family of concepts” (FR, p. 95). In fact, it is vague, not idiosyncratic.

Stich’s pet target among accounts of justified belief is the version of reliabilism presented in Goldman’s book. Unfortunately, he fails to defend his contention that this or any other developed form of reliabilism should be regarded as an *analysis* of the concept of justified belief. In fact, reliabilist theories attempt to spell out in substantive terms what it is in virtue of which a belief is justified, e.g., being the product of a reliable belief-forming process. What they take to be analytic, call it bare bones reliabilism, is that being justified is that property of beliefs in virtue of which they tend to be true. The idea here is that if this property, whatever it is, does not effect the “truth connection” (conduciveness to being

true), then it isn’t what underlies justification in the sense relevant to epistemology. This analytic thesis may be contentious, as analytic claims in philosophy often are, but it does not comprise the bulk of any reliabilist theory. So when Stich complains that reliabilist theories (and, for that matter, other theories in “analytic epistemology”) are parochial and explications merely of our “local concept,” he is misrepresenting their character. Besides, if it is merely our local concept that imposes this simple requirement of a truth connection, then Stich should provide at least a hint about the nature of the alternatives to our concept of epistemic justification that might prevail in other cultures.

Not only does Stich misrepresent reliabilist theories as conceptual analyses, he claims without argument that “these embellishments on bare bones reliabilism, though they may do a better job at capturing intuitions. will do a worse job at linking justification to truth” (FR, p. 98). Citing only the normal worlds approach given in Goldman’s book, Stich remarks, “if it’s truth we’re after, then we should exploit processes that are likely to produce truth in *our* world, not in *normal* worlds” (p. 98), and does not mention alternatives, including Goldman’s own revised theory, which attempt to do just that. In “Strong and Weak Justification”²⁰ Goldman distinguishes two conceptions of justified belief and explains that normal-worlds reliabilism applies only to weak justification. I developed a similar distinction in “A Rationale for Reliabilism,”²¹ but it was between justified belief and justified believers. The point of the latter notion, as with Goldman’s weak notion, is to allow for the case of a weird world, such as a demon world, where a justified believer is systematically mistaken, so that being justified does not necessarily conduce to having true beliefs. But the primary notion of justified belief (or Goldman’s strong justification) does involve the truth connection. These are but two versions of reliabilism which really are after truth in our world. Of course they do not offer advice on how to get it, so Stich’s comment that “if an analytic epistemological theory is taken to be part of the serious normative inquiry whose goal is to tell people which cognitive processes are good ones or which ones they should use, then for most people it will prove to be an irrelevant failure” (FR, p. 91) is itself irrelevant. Stich’s own advice, his pluralist pragmatist relativism, seems best suited for what might be called Pascal’s demon world, the world in which the beliefs to have are the ones you are best off having, regardless of their truth values. In that world, unfortunately, although there would be a certain interpretation function determining the property, TRUTH!!, such that the beliefs you are best off having are the TRUE!! ones, you could have no effective strategy for identifying the beliefs to have—unless you had *true* beliefs about the demon’s wishes.

So Stich has not made his case that justification and true belief are parochial notions of limited value. His assault on justification depends

on treating epistemological theories as analyses of the concept of justification, a local concept he claims to be of no special importance. In fact, this concept is distinguished by its connection to the concept of truth. It seems, then, that his assault on justification ultimately falls back upon his assault on truth. But that assault rests on the mistaken assumption that truth is determined by a certain interpretation function, one among many, and that this is the one that is “intuitively sanctioned.” The trouble is that there is no such function, for we don’t have intuitions about the truth conditions of our belief-brain-states. Besides, the ones for rivals to truth do not seem specifiable except in terms of *truth* conditions, so that it is hard to see how one could go about pursuing the TRUTH!. In any case, interpretation functions have nothing to do with the value of true beliefs, and, indeed, aiming at having true beliefs is not a matter of valuing truth.²²

Well, now that truth and justification have been defended, what about the American Way? Defending that would require upholding truths we hold to be self-evident. Perhaps that job is best left to the Man of Steel.

San Francisco State University
San Francisco, California

NOTES

¹ Cambridge, Massachusetts: Bradford/MIT 1990. Page references to the book will be to FR.

² “Stich Symposium,” *Philosophy and Phenomenological Research* 51 (1991): 197–213, which includes a précis of Chapters 4 and 5, commentary by L. Jonathan Cohen, Alvin I. Goldman, Gilbert Harman, and William G. Lycan, followed by Stich’s reply. Page reference to this symposium will be to SS.

³ Similar rephrasing is possible for Goldman’s observation that one “prefers (believes *p* and *p* is true) to (be agnostic about *p* and *p* is true) to (believe *p* and *p* is false)” (SS, p. 191).

⁴ Oxford: Blackwell 1990, p. 65. Horwich’s schema actually says “if and only if,” but this is too strong, since most of us are not so curious as to want to believe *every* truth.

⁵ See, for example, Paul A. Boghossian, “The Status of Content,” *Philosophical Review* 99 (1990): 157–184, the subsequent exchange between Michael Devitt (“Transcendentalism about Content”) and him (“The Status of Content Revisited”) in the *Pacific Philosophical Quarterly* 71 (1990): 247–263 and 264–278, and further comments (“Transcending Transcendentalism: A Response to Boghossian”) by Devitt and Georges Rey in the *Pacific Philosophical Quarterly* 72 (1991): 87–100. Their topic is the coherence of eliminativism about mental contents, and much of their debate presupposes a genuine distinction between two kinds of truth (or truth conditions), robust and deflationary. If Horwich is right, the distinction is untenable; indeed, he contends that there is no *property* of being true, hence no cost in giving it up.

⁶ Whereas Harman’s immanent truth “applies only to sentences within [one’s] own language of thought” (SS, p. 197), one of Horwich’s main themes (another is that most disputes involving the notion of truth are really about something other than truth) is that minimalist truth is applicable to languages other than one’s own, both actual and possible,

including ones with greater expressive power. If this point is correct, it provides a powerful rejoinder to Stich, when he says in reply to Harman, “those who think that having immanently true beliefs is intrinsically valuable are epistemic reactionaries [not merely conservatives]” (SS, p. 211).

⁷ At the object level it is very difficult even to state Stich’s position. If he claims that beliefs can do better than be true, what about the information states of systems that do not have beliefs, such as thermostats, guidance systems, simulations, and telephone books? Would Stich claim that information is an idiosyncratic notion and has lots of competition, such as schminformation?

⁸ Strictly speaking, the question is whether these sentences are true *simpliciter*, not true in a certain language. I think Tarski’s “true-in-L” leads to the misapprehension that there are as many kinds of truth as there are languages. Sentences may be language-bound, but that doesn’t make truth language-relative. So, for example, a better format for the T-sentence “‘Der Schnee ist weiss’ is true-in-German iff snow is white” would be “The German ‘Der Schnee ist weiss’ is true iff snow is white.” It seems to me that one source of the resistance to minimalism about truth is the supposition, due to Tarski, that there is not only the predicate “true” but predicates (as many as there are languages) of the form “true-in-L.”

⁹ It is limited in the sense of being a partial function—it fails to supply semantic interpretations for various sorts of “syntactically characterizable pattern of interaction among sentences” (FR, p. 113). However, Stich’s examples are so frivolous, involving what amount to numerological properties (they could matter only in a very far-fetched world), that it is doubtful that the interpretation function for any worthwhile TRUTH* would take them into account either. More to the point are states that are worthy of theoretical notice, ones that may be cognitively salient, such as non-propositional pictorial states and the uninterpretable, subconceptual states of connectionist models. But it seems that the interpretation function for TRUTH* would not take these into account either.

¹⁰ Stich boldly describes “the current intuitive theory of interpretation [as] the result of a long evolutionary culling process” (FR, p. 122), although “no one knows in any detail just how these intuitions arise” (FR, p. 120). He suggests that “in individuating the ‘constructions’ from which mental sentences are built, the pattern of causal interactions plays a certain role” (FR, p. 113), but he says nothing about how people could have cognitive access to mental sentences in these terms, much less have intuitions about their truth conditions.

¹¹ These are tendencies, not ironclad relations, for one can fail to draw a conclusion or refuse to draw a conclusion and give up a premise instead.

¹² Stich’s subsequent arguments are set against the background of a specifically “causal/functional version” of truth theory, based on Hartry Field’s account and “drawn from a variety of sources including . . . the Putnam-Kripke causal theory of reference and functionalism in the philosophy of mind” (FR, p. 106), but in the last section of Chapter 5 he generalizes all but one of his arguments to apply to any version of truth theory.

¹³ Perhaps Goldman had the latter problem in mind when he suggested that “TRUTH* and TRUTH** . . . seem to be complexes of I.F.s and truth values” (SS, p. 190).

¹⁴ He should say “additional schemes,” since for him a given sentence, and each of its referring constituents, has all the “semantic” properties given by the innumerable interpretation functions.

¹⁵ Stich imagines that we have an “intuitive strategy for mapping beliefs to propositions,” part of which “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” (SS, p. 182).

¹⁶ Even if one could identify it, believing its negation would be based not on epistemic grounds but, à la Pascal’s wager, for its practical benefits. Stich, of course, rejects any distinction between epistemic and practical reasons for belief. Goldman defends it (SS, pp. 192–193).

¹⁷ Stich questions Cohen's confidence that "it couldn't possibly be the case that some fallacious principle passes the reflective equilibrium test" (SS, p. 208), while Cohen complains that Stich's pragmatism precludes "assessing validity and utility independently" (SS, p. 188). Clearly each is begging the question against the other, but the question is moot if, as Harman suggests, the "exalted state" of reflective equilibrium is unattainable.

¹⁸ Cambridge, Massachusetts: Harvard University Press 1986.

¹⁹ These epistemic virtues concern having knowledge, not getting it. There are also such non-epistemic cognitive virtues as what Goldman (*op. cit.*, p. 26) calls "power" and "speed," which, despite their obvious pragmatic importance, Stich does not take up.

²⁰ *Philosophical Perspectives* 2 (1988): 51-69.

²¹ *The Monist* 68 (1985): 248-263.

²² I wish to thank Brian Loar for his perceptive comments on earlier drafts.

HOW NOT TO FLIP THE SWITCH WITH THE FLOODLIGHT: CAUSATIVE-INCHOATIVES, THE INSTRUMENTAL 'WITH', AND THE IDENTITY OF ACTIONS

BY

PATRICK FRANCKEN
and LAWRENCE BRIAN LOMBARD

I. *Causative-Inchoatives*

No action is a flipping of the switch unless the switch comes to be flipped, no action is an alerting of the prowler unless the prowler becomes alert, and no action is a closing of the door unless the door gets to be closed.¹ This is not an accident; it is a matter of meaning.

Terms such as 'the flipping of the switch', 'the alerting of the prowler', and 'the closing of the door' are, in fact, ambiguous. In one sense, they refer to events that the switch, the prowler, and door undergo, events that terminate, respectively, in the switch's being in the state of being flipped, the prowler's being alert, and the door's being in the state of being closed. In another sense, they refer to the actions and events that bring about those events. The flipping of the switch is, in this latter sense, just the action that, say, Smith, performed that brought about the switch's being in the flipped position. The rock's breaking of the window, in this sense, is just the event, whose subject was the rock, that resulted in the breaking of the window in the first sense, an event whose subject was the window and which consisted in the window's undergoing a change that terminated in its being in pieces.

It seems clear that, for at least a very large class of terms of this sort, the relation between the referent of such a term in its latter sense and the referent of that term in its former sense is *causal*. The rock's breaking of the window in the latter sense *caused* the breaking of the window in the former sense, that is, the window's breaking. And to say that Smith