Rational Credence and the Value of Truth

Allan Gibbard
University of Michigan
Ann Arbor, Michigan 48109
U. S. A.

Belief aims at truth—or so it is said, and there must be something right in the saying.[1] Belief, though, can't aim literally; it's *we* who aim. We aim, moreover, in acting—when we shoot an arrow, say, or more broadly, whenever we act to try to bring something about. Believing isn't an action; we can't believe at will. In some extended, metaphorical sense, perhaps, belief does aim at truth. Understanding this sense might even offer a philosopical key to belief. It might in particular tell us something important about rationality in belief. To employ the key, though, we'll have to understand what literally might underwrite this obscure dictum that belief aims at truth.

True belief is useful, we all know: armed with true beliefs, we can most effectively pursue our goals. A youth stands facing two doors; behind one, he has learned, is a lady to marry, and behind the other a ravenous tiger. If he knows the truth as to which door conceals the tiger, he can keep from being its prey. Otherwise, he takes his chances. True belief, then, has value as a guide to action, in pursuit of survival, wedlock, and a host of other goals that a person might have.

Sheer usefulness for such ulterior purposes, though, is not the only way that truth in belief can matter. We seek the truth, sometimes, purely to know it. Science at its purest is a search for important truths just to discover and have them. Indeed it's this disinterested search for truth, perhaps, that underlies rationality in belief of a special kind, a rationality that is not pragmatic but purely epistemic.

I share this last intuition, but I cannot make it work. There is such a thing as purely epistemic rationality, I accept, and it may sometimes contrast with pragmatic desirability in belief, with what it's rational to want to believe. (The man with indications that his wife is having an affair is a stock example; it may be epistemically rational for him to believe that she is, but rational for him to *want* not to believe it.) And epistemic rationality does

[1] Velleman, in "On the Aim of Belief" (2000), examines the claim "Belief aims at the truth," and offers a sense in which he thinks that it does. His interpretation of the dictum is quite different from the one I develop in this paper, though I think that his version and mine are compatible.

tie in closely with the value of truth as a guide to bringing about other things of value by our actions, such as wealth or happiness—with its "guidance value" as I'll call it. It has little to do, though, I am forced to conclude, with wanting truth for its own sake. Nothing about a goal of truth for its own sake leads to epistemic rationality as we know it.

1. Truth as the Aim

The "value of truth" is presumably the value of having true beliefs. It must also, it seems clear, include the disvalue of having beliefs that are false. Otherwise we could maximize our array of true beliefs just by believing everything whatsoever, true or false. It's these two kinds of value that are specifically epistemic, and that inform epistemic rationality. Belief in some sense aims at truth, but mostly, deliberate action isn't how we come to beliefs. If we are epistemically rational, we respond to our evidence directly with beliefs that are rational in light of that evidence. We don't will to believe with the aim of believing truly, as the youth might will to open the door on the left with the aim of finding the lady.

Here, though, is a way we might charitably interpret the claim that epistemically rational belief "aims at truth"—a way that will allow us to ask whether the aim of belief is truth for the sake of truth. Generally, to be sure, a person doesn't believe things by setting out to believe them. Still, we can try asking what it is to aim at anything in general, to have aims like avoiding tigers. If a person is epistemically rational, we can then hypothesize, then it is *as if* she chose her beliefs with the aim of believing truths and shunning falsehoods. She doesn't literally set out to believe truths, the way she might set out to get a high score on a test by intentionally putting down the right answers. But it is as if she did: it is as if she aimed at truth and away from falsehood in her beliefs in the same way one aims at any other goal.[2]

How, though, could this be? What the youth does with the aim of evading the tiger will depend on what he believes. If he believes the tiger lurks behind the door to the right,

[2] Velleman (2000) argues that a mental system doesn't count as forming beliefs unless its function is to produce true representations. He takes the dictum as covering all belief, not just belief that is fully rational. "To believe a proposition is to accept it with the aim of thereby accepting a truth" (p. 251). One's acceptance, that is to say, "is regulated, either by the subject's intentions or by some other mechanisms in ways designed to ensure that it is true" (p. 254). His term 'acceptance' covers not just beliefs but fantasies that motivate such things as children's make-believe, talking to oneself, and shaking one's fist at the man on the television screen. If the function in question is biological, then, as Nishi Shah argues, no normative consequences are strictly entailed ("How Truth Governs Belief", 2003, pp. 460–465). In this paper, I offer a normative interpretation of the dictum. My goal is not to interpret the dictum in the best way for all purposes, but to inquire whether we can explain epistemic rationality as somehow a matter of aiming at truth intrinsically. In adopting an interpretation in pursuit of this goal, I follow Joyce in many respects ("A Nonpragmatic Vindication", 1998).

he aims to avoid it by going to the left. Suppose, then, he wants to believe the truth and disbelieve all falsehood as to where the tiger lurks. If he believes that the tiger lurks to the right, he'll pursue his aim, if he can, by believing that it lurks to the right. This aim, it seems, has a strange property: no matter what he believes, it will be as if he had chosen his belief with the aim of believing the truth. He may of course lack all belief as to where the tiger lurks, but in that case, he would be at a loss for what to believe in pursuit of believing the truth, just as he is at a loss for which door to open in pursuit of evading the tiger. The aim of believing the truth, then, is empty in a way: if he thinks he knows what to believe in pursuit of it, that's what he already believes.

To aim at the truth is to guide one's beliefs by the evidence, we might try saying. This too, though, threatens to be empty. A person aims in light of what, rightly or wrongly, he *regards* as evidence. The youth aims to evade the tiger, and if he mistakes a datum for evidence as to where the tiger is, that's what he'll guide his actions by. What is it, then, we must ask, to regard a datum as evidence? Isn't it to adjust one's beliefs accordingly—or to perhaps to think that so adjusting them is warranted? If so, then to be responsive, by one's own lights, to the evidence may require no more than that one respond to data coherently, in a way that doesn't discredit itself. We'll need to understand what such coherence consists in, and if the injunction to heed the evidence helps in this, we need to discover how.

In another way, to be sure, the aim of truth is far from empty. The youth would take great efforts, if he could, to learn the truth of where the tiger lurks. Experiment, research, and deliberate observation are ways of acquiring true beliefs by seeking out further evidence. Often the costs are worth paying, and this shows that true belief is often of value, in a sense that has genuine import.[3] The value of truth as it is treated in this paper is highly relevant to questions of whether to bear a cost to acquire true beliefs, but the question I chiefly explore is a different one: Whether when we form our beliefs rationally, with no chance to seek out further evidence, it is as if we were forming beliefs at will with the aim of truth.

Grant, then, that we can't come to true beliefs by pulling on our own epistemic bootstraps, taking it as a real question what to believe in pursuit of truth. Aiming at truth in one's beliefs must have a strong element of circularity. Still, if the claim that belief aims at truth verges on being empty, isn't it at least true? With an epistemically rational person, it is as if, by her own lights, she were aiming at truth. This dictum might, indeed, offer a minimal test for epistemic rationality. A way of forming beliefs should at

[3] See Raiffa (1978, pp. 105–107 and chap. 7) on "buying" information, and Loewer (1993, 271–278) on the value of experiments. Horwich discusses the value of truth in *Truth* (1998); see pp. 44–46 on the instrumental value of truth and pp. 62–63 on its intrinsic value.

least satisfy this condition: if one forms beliefs that way, it will be as if one were, by one's own lights, forming beliefs voluntarily with the aim of believing truths and not falsehoods. Such a way of forming beliefs we might call *immodest*: it views itself as a way of acquiring truths and avoiding falsehoods. Immodesty is a minimal requirement, it might seem, on ways of forming beliefs. Absent it, after all, belief doesn't even aim at truth by its own lights. An empty virtue it may be, but indispensible.[4]

Immodesty will be my topic, and my puzzle will be this: If we look to belief for the sake of guidance, we'll indeed find that epistemically rational ways of forming beliefs are immodest. If instead, though, we look to truth's intrinsic value, to the pure scientific aim of truth for its own sake, immodesty eludes us.

All this will, of course, need elucidating. For epistemic rationality I'll assume a standard Bayesian account, and I'll draw on a background of Bayesian decision theory. Central to my argument will be theorems in a 1989 article by statistician Mark Schervish.[5])

2. Partial Credence

I have been taking belief, so far, as a matter of all or nothing. If the youth, though, formed some full belief as to where the tiger lurks, he would be rash. A full belief that the tiger is to the right might by luck be true, for all he knows, but then too it might be false. Rationality in belief often requires doubt; it demands some partial degree of credence. We need, then, to ask a further question: Suppose that an investigator, faced with her evidence, forms degrees of credence in a way that is epistemically rational. She doesn't, to be sure, form her partial beliefs at will, intentionally, in pursuit of a goal. Still, isn't it as if she did? Isn't it as if she somehow rationally chose her degrees of credence with the aim of truth in belief?

We are blessed with ample accounts of what it is to pursue a goal in light of one's degrees of credence. (I'll follow David Lewis and speak of a degree of credence, for short, as a *credence.*) The classic derivations in this vein stem from Ramsey and Savage, with Hammond's perhaps the most general and complete.[6] I don't in this paper try to reassess the arguments for this classical account of the rational pursuit of goals under uncertainty. The account is controversial, and I won't comment on the points in dispute. What I'll ask is this: Suppose the classic account of expected utility maximization does indeed describe

---

[4] The term "immodest" and the thought behind it are drawn from David Lewis, "Immodest Inductive Methods" (1971).

[5] Schervish, "A General Method" (1989), esp. Theorem 4.2, p. 1861.

[6] Ramsey, "Truth and Probability" (1931); Savage, *Foundations of Statistics* (1954); Hammond, "Consequentialist Foundations" (1988).

the rational pursuit of goals under uncertainty. Then can we say that belief aims at truth? I interpret this as before: When a person forms her credences with epistemic rationality, is it *as if* she were choosing her credences voluntarily, rationally aiming, in light of her credences, at truth in those very credences? Is an epistemically rational way of forming credences "immodest" in this sense?

More needs to be said to give a clear sense to this question. A complex of goals, on the classic account, is represented by a utility function. This function summarizes how one trades off one goal against another when one can't reach all of them with certainty. When credence is partial, a matter of degree, what utility function, we can ask, would characterize having the pure goal of truth in one's beliefs? Full credence in truths is sought and full credence in falsehoods shunned, to be sure—but ordinarily, one doesn't fully know how to achieve this aim. What of intermediate degrees of credence? In the case of a truth, if truth is what the agent seeks in her credences, then in her estimation, it seems, the higher the credence the better. The closer to full belief a credence is, we could say, the more "accurate" it is, and what the agent seeks is, in this sense, credences that are accurate.[7] For a falsehood, correspondingly, the lower the credence the better; the closer it is to a fully accurate credence of zero.

This is what we can draw from our intuitive concept of purely seeking truth in one's credences. Suppose, in particular, that only one claim is at issue: say, the claim that modern Europeans descend at least in part from the Neanderthals. (Call this claim S.) Take an investigator who values nothing but the accuracy of her degree of credence in this claim. The states which have greater or lesser utility in her eyes have two components: whether the claim is true or false, and her credence in the claim. We could represent her utilities, then, with a function $u(\mathbf{v}, x)$, where $\mathbf{v}$ is a truth value $\mathbf{t}$ or $\mathbf{f}$, and x is a possible credence, a real number with $0 \leq x \leq 1$. More conveniently, we can use two separate functions. Let $g_1(x) = u(\mathbf{t}, x)$, the value, in the believer's eyes, of having a credence of x in claim S where S turns out to be true. Let $g_0(x) = u(\mathbf{f}, x)$, the value, in her eyes, of having a credence of x in claim S where S turns out to be false. That she values truth and truth alone in her credence in S, then, seems to consist in satisfying this condition:

Condition $\mathcal{T}$: Function $g_1(x)$ increases strictly monotonically with x, and function $g_0(x)$ decreases strictly monotonically with x.

[7] Joyce, "A Nonpragmatic Vindication" (1997), pp. 586ff., speaks of the "gradational accuracy" of a degree of credence. Maher ("Joyce's Argument", 2002, 79) criticizes Joyce's "Norm of Gradational Accuracy", in part because it requires a standard of accuracy. Here I simply require that one who "aims at truth" maximize expected accuracy on some specification or other of the standard. Maher advocates looking to a concern to accept truths and not falsehoods for the tie of credence to truth. On this score I have sided with Joyce, but I won't pursue the debate further.

In other words, she thinks that in case S is true, the higher her credence in S the better, and that in case S is false, the lower her credence in S the better.

This allows, though, that there are many somewhat different ways a person could value truth in her credences. She could equally value every 1% increase in her credence in a truth and every 1% decrease in her credence in a falsehood. Alternatively, for a truth, she could be especially concerned with approaching full certainty: she might value the difference between 89% certainty and 99% certainty far more than that between 50% and 60% certainty. These differences will be represented by the functions g_1 and g_0: in the first case the functions are both linear, in that for some constants $k > 0$, a_1, and a_0,

$$g_1(x) = a_1 + kx; \qquad g_0(x) = a_0 - kx.$$

In the second case, we have

$$g_1(.99) - g_1(.89) \gg g_1(.60) - g_1(.50).$$

Either such function, it seems, fits the intuitive content we have teased out of the notion of exclusive concern with the truth. They indicate different emphases that such a concern might adopt—an emphasis, say, on certainty verses its lack, as opposed to equal concern with each possible 1% difference in credence.

Return now to our hypothesis that "belief aims at truth." With all or nothing belief, the hypothesis threatened to be empty; what happens now? When a person forms her credences with epistemic rationality, our hypothesis will now run, it is as if she were voluntarily chosing her credences with the pure aim of truth—that is to say, to maximize the expected accuracy or her credence. Accuracy we characterize with two functions g_1 and g_0 that satisfy Condition $\mathcal{T}$.

I begin with the good news: At least one variant of "aiming at truth" does fit this hypothesis. This result is well known; the utility function in question is the "Brier score" explored by George Brier in 1950. Later, though, comes more distrubing news: most variants of "aiming at truth" fail to fit this hypothesis. What, we can then ask, characterizes the subclass of variants that do fit it? A theorem of Schervish gives the answer: the utility functions that fit the hypothesis turn out to be the ones that could constitute a measure of *guidance value:* the value of an array of credences as a guide to choice in pursuit of other values—money, love, or succoring humanity, say. Intrinsic concern with truth, so far as I can discover, has nothing to do with what makes this subclass special.

3. Credence-Eliciting Functions

The rational believer of our inquiry has an array of credences formed with epistemic rationality on the basis of evidence. We ask whether those are the credences that, in light of her evidence, she most prefers to have, the ones that she would choose if she could choose her credences at will. Her only concern is, in some sense, how accurate her credences are; the emphases she places on various different aspects of accuracy are indicated by our two functions g_1 and g_0.

Another question of the same mathematical form has been much investigated. (I follow Joyce in pursuing the parallel.[8]) Beginning with Brier, a number of writers have asked what incentives would elicit an expert's true credences—a weather forecaster's credences, say, in rain for tomorrow.[9] The resulting payment to the expert, the idea is, can serve as an index of the quality or success of the expert's reported credences, how "close to the truth" those credences come. To aim to maximize such an index, we can say, would be to aim at truth in one's beliefs.

We engage as an expert informant, then, a *homo economicus* who has no intrinsic concern with reporting to us honestly. What incentives can we give him to reveal to us his genuine credences? What we pay him will depend on the credences he reports and the actual truth of the matter. Where x is his reported credence in a contingency S (rain tomorrow, say), we pay him (in utility) $g_1(x)$ if S turns out to be true and $g_0(x)$ if S turns out to be false. Since this takes the same mathematical form as our own problem of what sort of intrinsic concern with the truth would make one want the credences one has, I'll say that our problem is to characterize pairs of functions g_1 and g_0 that are *credence-eliciting.*

The Brier score is given by the following credence-eliciting pair of functions. For convenience, define $\bar{x} = 1 - x$. Then we let

$$g_1(x) = 1 - \bar{x}^2; \qquad g_0(x) = 1 - x^2.$$

(Define the believer's *inaccuracy* as the distance of her credence from full accuracy. Complete accuracy would be credence 0 for S false and 1 for S true; the inaccuracy of a credence in S is thus x for S false and $\bar{x}$ for S true. The Brier score penalizes the believer by the square of her inaccuracy.)

We can now ask whether, if a rational believer maximizes the expected value of her Brier score, she will choose, in prospect, the credences she already has. Let her actual credence $\rho(S)$ in S be α. What credence x would she choose to have, if she could? Her expected Brier score is

[8] Joyce, "A Nonpragmatic Vindication" (1998).

[9] Brier, "Verification" (1950) and others noted below.

$$\alpha(1-\bar{x}^2)+\bar{\alpha}(1-x^2)$$

As Brier showed, this is maximized when $x = \alpha$, so that one most prefers the credence one has.[10]

Consider, though, a different pair of functions g_1 and g_0 that satisfy Condition $\mathcal{T}$ and so indicate a way in which one might be concerned to to minimize a gauge of the the inaccuracy of one's credence. Suppose I aim to minimize my expected inaccuracy, rather than its square, so that for me,

$$g_1(x) = -\bar{x}; \qquad g_0(x) = -x.$$

My expected utility for having credence x, then, is

$$\alpha(-\bar{x}) + \bar{\alpha}(-x).$$

I maximize this by making my beliefs extreme in their certitude: if $\alpha > 1/2$, then my best bet is setting $x = 1$, and if $\alpha < 1/2$, then my best bet is setting $x = 0$. If my intrinsic concern with truth took this form, I would rationally advance this concern, if I could, by moving to an epistemically rash certitude, by jumping to whichever conclusion I found even slightly more plausible.

We can ask, then, what form an exclusive, intrinsic concern of truth must take for epistemic rationality to be prospectively the best policy. Equivalently, what functions are credence-eliciting? Here is the answer in a form that ties in with lessons about "guidance value" that I draw later. Let h be any smooth, increasing function of x. Then we can let g_1 and g_0 be functions that satisfy these conditions:

$$g_1'(x) = \bar{x}h'(x), \tag{1_1}$$

$$g_0'(x) = -xh'(x). \tag{1_0}$$

Smooth functions g_1 and g_0 are credence-eliciting, the theorem is, just in case there exists some smooth, increasing function h such that (1_1) and (1_0) obtain.[11]

[10] Her expected utility is $u(x) = 1 - \alpha\bar{x}^2 - \bar{\alpha}x^2$. The first derivative is $2\alpha\bar{x} - 2\bar{\alpha}x = 0$ for $u(x)$ maximal, which holds just for $x = \alpha$.

[11] This is a variant of Theorem 1 of Shuford, Albert, and Massengill, "Admissible Probability" (1966), p. 128. Savage ("Elicitation", 1971) gives the answer in a different form. Schervish ("A General Method", 1989, p. 1863b) puts the result in terms of a measure λ rather than an increasing function h. He relaxes the restriction to smooth functions in his Apperdix, pp. 1874–78.
Schervish and others speak of "proper" and "strictly proper" scoring rules. The rule is *proper* iff it makes honest reporting of credences optimal, and *strictly proper* iff it make honest reporting uniquely optimal. By my term 'credence-eliciting' I mean strictly proper.

The function h turns out to be of special significance: its slope indicates the urgency the believer ascribes to getting credences right, by her lights, in the vicinity of x. If, for instance, the believer is especially concerned with distinguishing degrees of near certainty, then the slope of function h will be especially great for credences near one. That's a story for later, though. For now, the two points to note are these. First, many different, non-equivalent payment schemes turn out to be credence-eliciting. Many different balances of concern turn out to be characterized, for some smooth increasing function h, by equations (1_1) and (1_0), and thus to engender immodesty. Here are a few examples:

$h(x)$	$g_1(x)$	$g_0(x)$
$2x$	$x(2-x)$	$-x^2$
$3x^2$	$x^2(3-2x)$	$-2x^3$
$-ln(1-x)$	x	$x + ln(1-x)$

Still, the requirement of immodesty is quite constrictive. The functions g_1 and g_0 must bear a tight relation to one another. We can let g_1 be any smooth increasing function whatsoever, but once g_1 is chosen, g_0 is determined apart from a constant. Likewise, g_0 can be any decreasing function, but it determines g_1 down to an arbitrary constant. Specify that $g_1 = -\bar{x}^2$, and we must have $g_0 = K - x^2$, the Brier rule plus a constant. Specify that $g_1 = x$, and we must have $g_0 = K + x + ln(1-x)$. The vast bulk of pairs g_1 and g_0 won't be credence-eliciting; indeed those that are will be an infinitesimal fraction of all the possible pairs of smooth increasing functions.

Explicitly (though these precise formulas won't matter for the rest of what I have to say), we require that for every x,

$$\bar{x}g_0'(x) = -xg_1'(x), \tag{2}$$

so that given function g_1, we have

$$g_0(x) = K - \int_0^x \frac{z}{\bar{z}}\, g_1'(z), \tag{3}$$

Only quite a special relation between the functions g_1 and g_0, then, allows epistemic rationality to be the best policy, by one's own rational lights, in pursuit of truth. We need a term for pairs of functions g_1 and g_0 that satisfy condition (2) and hence formula (3). We must ask what makes for this special relation. In the meantime, I'll coin a label for it: a pair of functions g_1, g_0 that satisfies (3) I'll call *SAM-qualifying,* after the authors who discovered the relation.[12]

[12] Conditions (2) and (3) are variants of formulas in Shuford *et al.*, "Admissible Probability" (1966), 128–129.

4. Credence as Guide to Action

An intrinsic concern for accuracy in one's credences is not the only reason one might care what one's credences are. As I stressed at the outset, credence guides action in pursuit of any array of aims whatsoever—avoiding tigers and gaining a lady, for instance. The value of one's credences in this role I am calling their *guidance value.* We can ask what array of credences will, by a rational person's lights, have the maximal expectation of guidance value. The answer turns out to be, whatever credences one in fact has. So long as they are formally coherent—satisfy the standard conditions on probabilities—any array of credences whatever will be, in this new sense, "immodest".

Previously, I characterized immodesty as follows. An epistemic policy is immodest just in case when a person forms credences by that policy, it is as if she were forming them voluntarily, rationally aiming, by her lights, at truth in her credences. This required a specification of what it is to aim at truth in one's credences. The specification must take the form, I said, of a pair of SAM-qualifying functions g_1 and g_0, functions related to each other as in (2). We can speak, then, of an epistemic policy P as *truth-immodest* with respect to functions g_1 and g_0 just in case, as calculated according to policy P and determined by g_1 and g_0, the expected payoff of adopting policy P is at least as great as that of adopting any alternative epistemic policy. With respect to some pairs of functions that specify a way of aiming at truth—the ones that are SAM-qualifying—all coherent epistemic policies are truth-immodest. With respect to other pairs that equally well specify a standard of aiming at truth, however, no epistemic policy will be truth-immodest. We now can specify a similar but distinct notion of immodesty, which I'll call being "guidance-immodest". Roughly, an epistemic policy P will be guidance-immodest just in case, as calculated in accordance with policy P, policy P has maximal prospective guidance value.

The guidance value of an epistemic policy, though, depends on a number of things: the choices one must make, the facts that determine the outcomes of those choices, and the values of those possible outcomes. We need to spell out more fully, then, what constitutes guidance-immodesty in an epistemic policy, and with respect to what.

For truth-immodesty, I hypothesized at the outset that, with an epistemically rational thinker, it is as if she had chosen her credences voluntarily, with an intrinsic aim of accuracy. That turned out to be true only if the concern with accuracy takes a form that is sharply constrained, only if it is SAM-qualifying. Now, though, we can try out a different hypothesis, one that invokes guidance-immodesty. With an epistemically rational thinker, it will turn out, it is as if she had chosen her credences voluntarily and entirely for their prospective guidance value. It is as if she had chosen them for their value in guiding an array of choices she might be faced with, in pursuit of ultimate goals that don't include accuracy in her credences.

Guidance value, of course, is far from the only kind of value one's credences can have. Beliefs can be comforting. They can be empowering. They can link one to others in a fellowship of conviction. I'll label all the kinds of value that credences can have apart from their guidance value as *side value.* The import of the Schervish theorem I'll be presenting is this: Suppose one's credences have no side value whatsoever but only guidance value. Then epistemically rational credences are prospectively the best ones to have.

With "side value" defined just as any value apart from guidance value, I should specify more clearly what "guidance value" means. The rough idea is clear, I hope: The gallant youth opens whatever door he more strongly believes conceals the lady. His credences have high guidance value if they lead him to the lady, and terrible guidance value if they lead him to the tiger. He thus brings about a result—that he finds the lady or that he finds the tiger—acting voluntarily, on a policy that directs what to do as a function of his credences.[13] He maximizes expected utility, as calculated in the standard way using his credences and some scale of valuation. We gauge the outcome using that same scale of valuation. With respect to that scale of valuation, then, the guidance value of an array of credences ρ is the value, on that scale, of all that he would bring about by performing, at will, whatever acts have highest expected utility as reckoned using credences ρ.

More generally, take any policy F for acting on the basis of one's evidence, and consider a scale of value v. The guidance value of F with respect to scale v is the value, on scale v, of all that would be brought about by the voluntary acts directed by policy F. Now a coherent policy F for action, I am assuming, will consist in a coherent epistemic policy F_E and a policy of maximizing expected utility as reckoned with some scale of value v and the credences ρ yielded by policy F_E. (What the policy directs in case of ties won't matter for our purposes, as long as it directs some act that maximizes expected utility so calculated.) We can speak, then, of the guidance value of credences σ with respect to scale of value v. This is the value, on scale v, of all that would be brought about by the voluntary acts maximizing expected utility calculated with credences σ and value scale v.

The actual guidance value of one's credences, the guidance value they in fact turn out to have had, is thus a matter not only of what choices one would rationally make, in light of those credences and in pursuit of one's aims, but also of how things turn out to

[13] An act "brings about" its causal consequences, but not only those. For a driver, for instance, the act of pushing the turn signal lever down may bring it about that one thereby signals a left turn. That one signals a left turn isn't, strictly, caused by the act of pushing the lever, as the blinking of the signal light is caused by that act, but the act does bring it about that one signals the turn. If the signaling itself has value, that counts toward the utility of pushing the lever. See Kim, "Noncausal Connections" (1974).

be. Suppose that in fact, the tiger is behind the door to the right, and the youth cares, rationally, only about escaping the tiger. Rationally, then, he will go left—and so escape the tiger—just in case his credence in the tiger's being to the right is greater then $1/2$.[14] Let the unit of value be the ramsey or "ram", and let the value of escaping the tiger be 100ram. Then as matters stand, any greater-than-even credence that the tiger is to the right has 100ram greater guidance value than does any less-than-even credence.

Prospective guidance value, then, in light of one's credences, is one's subjective expectation of this actual guidance value. The youth, suppose, for some reason or other, rationally has a credence of 60% that the tiger is to the right. Then the prospective guidance value of any degree of credence $x > 1/2$, as opposed to any degree $y < 1/2$, by his lights, will be $.6 \cdot 100\text{ram} + .4 \cdot 0\text{ram} = 60\text{ram}$. Prospective guidance value, by one's lights, thus depends on one's credences. Even if we could choose what credences to have, we couldn't bootstrap our way to credences prospectively useful as guides to choice by starting with a mind empty of all credence. We can ask, though, whether a way of forming credences is *immodest* from the standpoint of guidance value: whether, if one forms one's credences that way, one will attribute to credences formed in that very way, then, a maximal expectation of guidance value.

The actual guidance value of an array of credences depends on the choices one encounters. Likewise, their prospective guidance value, by one's own lights, depends on one's credences as to what choices one will encounter. Think of life as a single big and complex gamble, and suppose one is uncertain what complex gamble one will face. I'll consider only the special case where a single contingency is to be gambled on, but at odds one doesn't know in advance. Set up, then, a continuum of bets on a claim S. One will take or reject each of these bets, suppose, on the basis of a credence x in S that one choses at will.

For each β with $0 < \beta < 1$, then, set up an infinitesimal bet G_β at odds $\bar{\beta} : \beta$. Being guided by a credence $\rho(S) = x$ in S consists in the following: accepting all bets G_β for $\beta > x$ and rejecting all bets G_β for $\beta < x$. It turns out that one's most favorable prospect, by one's own lights, will be to choose as one's guiding credence x one's true credence $\rho(S) = \alpha$.[15]

Schervish's result (in my own words and apart from some niceties) is this:[16]

---

[14] I'll ignore, in this discussion, points of indifference, as when his credence is exactly $1/2$.

[15] It won't then matter, prospectively by one's own lights, how we decide on bets G_β with $\beta = x$, since one will be prospectively indifferent between accepting and rejecting a bet G_β when one's credence in S is β.

[16] Schervish, "A General Method" (1989). This is essentially his Theorem 4.2, p. 1861, which concerns scoring a set of actual forecasts. As he later notes, "The results of the previous sections can be easily translated to results concerning the expected score" (p. 1869 [h]). Where f is the average loss to the decision maker

THEOREM: Smooth functions g_1 and g_0 are credence-eliciting if and only if for some possible continuum of bet offers and a policy of accepting any bet offer G_γ exactly when $\gamma < x$, $g_1(x)$ gives the expected payoff of the policy given S, and $g_0(x)$ gives the expected payoff of the policy given $\bar{S}$.

As Schervish summarizes this finding, "Scoring rules are just a way of averaging all simple two-decision problems into a single, more complicated, decision problem" (1873c).

Schervish also gives us an interpretation for the function h, a function which was left an uninterpreted mystery in our previous discussion. The slope $h'(x)$ of h at x turns out to be the payoff density of bets G_x at odds $\bar{x} : x$. It is the payoff density at those odds in the package of bets that one accepts or rejects on the basis of one's chosen guiding credence x. This will explain how different functions playing the role of h represent different ways of aiming at the truth. For any given x, the slope $h'(x)$ indicates the urgency of probability discriminations in the vicinity of x. If α is the best credence to report, then to a second approximation, a "mistake" of size of size δ costs, prospectively, $\frac{1}{2}h'(\alpha)\delta^2$. The slope of h also indicates, to a second approximation, the value of acquiring a piece of evidence as to whether S obtains. Appendix I demonstrates all this.

The choice among guidance-immodest pairs of functiona g_1, g_0 affects not which credence you will regard as optimal, but your preference among those you regard as non-optimal. Different "mistakes" in choosing your credences would affect different choices you might make, and the question is how these choices prospectively matter.[17] Whatever kinds of prospective choices you emphasize in your concern for truth, whatever functions g_1 and g_0 indicate these emphases, the following will obtain: so long as you coherently value truth solely for the sake of guidance, you do prospectively best, by your own lights, with the credences you have. In this regard, the choice among guidance-immodest specifications of the aim of truth does not matter.

5. Discussion: Seeking Truth

Epistemically rational belief aims at the truth—or so we might think. My hypothesis at the outset, recall, was this: When a subject forms her credences with purely epistemic rationality, it is as if she chose her credences voluntarily with the pure aim of accuracy. I assumed that a necessary condition for epistemic rationality is *coherence*, in the sense of

and a "proper" scoring rule is one that is, in my terms, credence-eliciting, Schervish writes, "Integrating f is essentially equivalent to calculating a proper scoring rule" 1862a.

[17] Schervish explains why different scoring rules order forecasters differently (1989, p. 1862e): one measure has more of its mass at probabilities where the one forecaster does better, and the other measure, where the other forecaster does better. (His measure λ is the one induced by my function h.)

satisfying the standard axioms of probability, and that a necessary condition for rationally aiming at something is that one maximize expected value, on some scale of value that specifies the aim, as calculated using some coherent array of credences. The emphases the subject gives to different aspects of accuracy in her credences are represented by the functions g_1 and g_0, the first increasing and the second decreasing. The hypothesis holds true, it turns out, just in case the functions g_1 and g_2 are SAM-qualifying—just in case they bear a highly constrained relation to one another, the relation given by (2) and by (3). If they are SAM-qualifying, then the hypothesis is fairly empty: any way of forming coherent credences will fit the hypothesis. (What happens if credences are not coherent I have not here investigated.[18])

More explicitly put, our finding was this: Suppose a subject aims purely at accuracy in her credence x for claim S, in that her utility is given by an increasing function $g_1(x)$ for the case of S true and a decreasing function $g_0(x)$ for the case of S false. Then if she forms her credences with epistemic rationality, it is as if she chose her credences voluntarily with a pure, SAM-qualifying aim of accuracy. It is not, however, as if she chose them with a aim of accuracy that fails to be SAM-qualifying.

An aim of accuracy is SAM-qualifying in this sense, we have seen, just in case it exactly matches a pure concern with guidance value, given some possible, sufficiently rich prospect for what "bets" one will face in life. One could value accuracy in one's credences intrinsically and prefer the epistemically rational credences one has—but only if one's valuations of truth takes this special form, only if it is SAM-qualifying.

So could a pure concern with truth for its own sake explain epistemic rationality? It seems not. In the first place, of course, as with any aim, to pursue an aim of accuracy, one must already have credences in place—and to pursue such an aim rationally, one must have rational credences already in place. In the second place, though, even then, a concern with accuracy leads one to prefer rational credence only if the concern is SAM-qualifying, only if it takes a special form. Simply wanting truth or accuracy for its own sake does not explain this form. Wanting truth entirely for the sake of guidance *would* explain it—and this is the only explanation we have found.

Might there be some further explanation? Perhaps we have not teased out all the requirements that are implicit in the ordinary notion of a "pure concern with truth" in one's beliefs. Joyce places two further requirements on a pure concern for accuracy, which he calls Normality and Symmetry (1998, p. 596). With these conditions, he is able to prove a strong result—one that it would be good to prove for the framework of this paper. All and only coherent arrays of credence, he shows, satisfy the dictum that belief aims

[18] Joyce's theorem (1998) gives the answer for a class of functions g_1 and g_0 that include the Brier score.

at truth. The dictum he interprets as I have have interpreted it in this paper, but with Joyce's two additional conditions in place. These conditions, though, have the effect of restricting functions g_1 and g_0 to the Brier rule and scoring rules closely related to it. In my terms, they require that that the function h be linear, that its slope h' be constant. Most of the scoring rules I gave as examples violate this condition. A constant slope for h means, as Schervish shows, that the subject has equal concern for every 1% difference in credence. I argued early on in this paper that this is no requirement for a pure concern with truth, that an investigator's pure concern with truth is unsullied if he cares much more about getting near to certainty than about fine differences in middling credences.[19]

Where does this leave pure concern with truth? I find the matter puzzling. Earlier, I distinguished the guidance value of a credence from its "side value". A side value, I said, is any value a credence might have apart from its pure guidance value: the comfort, empowerment, or fellowship it brings, for example. Now by this strict definition, any intrinsic value that truth might have counts as a side value. As with other side values, we have seen, an intrinsic concern for truth, for some form of accuracy in one's credences, can lead one to prefer credences that are not coherent, and so not epistemically rational.

Intrinsically valuing truth, though, won't do this if one's concern with accuracy takes the right restricted form—if it is "qualifying", if it matches a guidance value that one's credences might have had. And it does seem that we all have some intrinsic concern with the truth. We're all curious, after all. It does seem too that pure science purifies and elaborates this concern. That won't lead to trouble if the aim of truth takes a form that is qualifying, as I have defined the term, if it mimics possible guidance value.

An intrinsic concern for the truth can't be what we demand of pure investigators unless it mimics a concern for truth for its guidance value. An intrinsic concern for truth, though, so long as it does match aiming at truth purely for its guidance value, seems to be just what characterizes rational curiosity and the purest of science. But concern with guidance value is instrumental, not intrinsic. Why must the pure aims of science mimic it? I don't know the answer.

Does belief, then, aim at truth? Yes, but in a special way. Belief, we have seen, aims at truth, but not perhaps for the sake of truth itself. Belief aims at truth for the sake of guidance.[20]

---

[19] The argument was in Section 1, when I introduced Condition $\mathcal{T}$. Maher criticizes Symmetry ("Joyce's Argument, 2002, 76–78).

[20] Earlier versions of this paper were presented to the Center for the Philosophy of Science at the University of Pittsburgh and to the Creighton Club meeting at Cornell University. I am grateful for discussion on both occasions. James Joyce has been of especial value to me in discussing this paper and in guiding me into its subject and the literature of the field. Teddy Seidenfeld pointed me to his "Calibration" (1985),

Appendix I. Function h as an Indicator of Urgency.

What is the prospective cost to our advisor of reporting the wrong credence. Her true credence is α, suppose; what would she lose by guiding herself by a credence $x = \alpha + \delta$ (where δ may be positive or negative)? To a first approximation, small misreports are prospectively costless: only a small range of bet offers get decided differently, and the mistake she makes on any one of those is prospectively small. Where $\mathcal{E}(x)$ is the expected payoff of adopting x as one's guiding credence and we define

$$\Delta\mathcal{E}(\alpha) = \mathcal{E}(\alpha + \delta) - \mathcal{E}(\alpha),$$

then to a second approximation we have

$$\Delta\mathcal{E}(\alpha) = \mathcal{E}'(\alpha)\delta + \tfrac{1}{2}\mathcal{E}''(\alpha)\delta^2 = \tfrac{1}{2}\mathcal{E}''(\alpha)\delta^2,$$

since $\mathcal{E}'(\alpha) = 0$. Since we have

$$\mathcal{E}(x) = \alpha g_1(x) + \bar{\alpha} g_0(x)$$

the first derivative is

$$\begin{aligned} \mathcal{E}'(x) &= \alpha g_1'(x) + \bar{\alpha} g_0'(x) \\ &= \alpha\bar{x}h'(x) - \bar{\alpha}xh'(x). \end{aligned} \tag{4}$$

The second derivative, then, is

$$\mathcal{E}''(x) = (\alpha\bar{x} - \bar{\alpha}x)h''(x) - (\alpha + \bar{\alpha})h'(x).$$

Thus $\mathcal{E}'(\alpha) = 0$, as we already know, and $\mathcal{E}''(\alpha) = -h'(\alpha)$. The slope of function h at α, then, indicates how urgent it is to avoid mistaken reports in the vicinity of α. To a second approximation, if α is the best credence to adopt as a guide, then a mistake of size δ prospectively costs $\frac{1}{2}h'(\alpha)\delta^2$.

Now consider buying information. In one sense, one values truth in one's belief concerning a proposition T if one is willing to pay a cost to learn whether or not S obtains, or to acquire evidence as to whether or not S obtains. For our advisor who has credence α in S and faces payoffs functions g_1 and g_0, the prospective value of learning the truth about S is

---

276–277, which led me to Schervish's paper. I am grateful to Aaron Bronfman, Paul Horwich, and David Velleman for helpful Discussion.

$$\alpha[g_1(1) - g_1(\alpha)] + \bar{\alpha}[g_0(0) - g_0(\alpha)].$$

For $g_1(1) - g_1(\alpha)$ is her gain from reporting a credence of one instead of α if S obtains, and $g_0(0) - g_0(\alpha)$ is her gain from reporting a credence of zero instead of α if S doesn't obtain. These are both positive, since g_1 increases and g_0 decreases, and so learning whether or not S is true has a positive prospective guidance value. (Of course its actual guidance value may turn out negative; even for the idealy rational, a little truth can be a dangerous thing.)

This leads to an analysis of how the payoff functions g_1 and g_0 induce her to value evidence as to whether S obtains. For sufficiently weak evidence, the slope of the function h indicates the value of the evidence to her, to a second approximation. Suppose that the question whether T bears on the question whether S, and that this bearing exhausts the guidance value of T. Let her credence in T be β, and suppose learning that T would shift her credence in S by an amount δ_1, whereas learning that T is false would shift her credence in S by an amount δ_0. Then we have seen that to a second approximation, her prospective gain from learning that T obtains, in case it does, is $\frac{1}{2}h'(\alpha)\delta_1^2$. Her prospective gain from disbelieving T, in case T doesn't obtain, is $\frac{1}{2}h'(\alpha)\delta_0^2$. Her prospective gain from learning the truth as to whether T, then, is

$$\begin{aligned} &\tfrac{1}{2}\beta h'(\alpha)\delta_1^2 + \tfrac{1}{2}\bar{\beta} h'(\alpha)\delta_0^2 \\ &\quad = \tfrac{1}{2}h'(\alpha)[\beta\delta_1^2 + \bar{\beta}\delta_0^2] \end{aligned}$$

References

Brier, Glenn W. (1950). "Verification of Forecasts Expressed in Terms of Probability". *Monthly Weather Review* **78**:1 (January), 1–3.

Hammond, Peter (1988). "Consequentialist Foundations for Expected Utility". *Theory and Decision* **25**, 25–78.

Horwich, Paul (1998). *Truth*, 2nd edition (Oxford: Oxford University Press).

Joyce, James (1998). "A Nonpragmatic Vindication of Probabilism". *Philosophy of Science* **65**, 575-603.

Kim, Jaegwon (1974). "Noncausal Connections." *Nous* **8**, 41–52.

Lewis, David (1971). "Immodest Inductive Methods". *Philosophy of Science* **38**, 54–63.

Loewer, Barry (1993). "The Value of Truth". E. Villanueva, ed., *Philosophical Issues* **4**: 265–280 (Atascadero, CA: Ridgeview Publishing Company).

Maher, Patrick (2002). "Joyce's Argument for Probabilism". *Philosophy of Science* **67**, 73–81.

Raiffa, Howard (1968), *Decision Analysis* (Reading, Ma.: Addison Wesley).

Ramsey, Frank Plumpton (1931), "Truth and Probability". *The Foundations of Mathematics and other Logical Essays* (London: Routledge & Kegan Paul).

Savage, Leonard J. (1954), *The Foundations of Statistics* (New York: Wiley).

Savage, Leonard J. (1971). "Elicitation of Personal Probabilities and Expectations". *Journal of the American Statistical Association* **66**, 783–801.

Savage, Leonard J. (1972), *The Foundations of Statistics* Second Edition (New York: Dover).

Schervish, Mark J. (1989). "A General Method for Comparing Probability Assessors". *Annals of Statistics* **17**:4, 1856–1879.

Seidenfeld, Teddy (1985). "Calibration, Coherence, and Scoring Rules". *Philosophy of Science* **52**, 274–294.

Shah, Nishi (2003). "How Truth Governs Belief." *Philosophical Review* **112**:3, October, 447–482.

Shuford, E.H., Albert, A. and Massengill, H.E. (1966). "Admissible Probability Measurement Procedures". *Psychometrika* **31**, 125–145.

Velleman, David (2000). "On the Aim of Belief". *The Possibility of Practical Reason* (Oxford: Clarendon Press), 244–281.