BARRY LOEWER

THE TRUTH PAYS*

Why is truth valuable? Why are true beliefs generally preferable to false beliefs and why should we often be willing to expend energy and resources to obtain the truth? Pragmatist theories of truth, whatever their shortcomings, are the only ones which attempt to answer these questions. According to James' version of the pragmatic theory:

> The possession of truth, so far from being an end in itself is only a preliminary means toward other vital satisfactions.... True ideas would never have been singled out as such, would never have acquired a class name, least of all a name suggesting value, unless they had been useful from the outset in this way.... Primarily, and on the common sense level, the truth of a state of mind means this function of *leading that is worthwhile*.... Our account of truth is an account of truths in the plural, of processes of leading; realized in *rebus*, and having only this quality in common, that they *pay*.[1]

James' view is that there is an intimate connection between true belief and success in practical action. Sometimes he writes as if he holds that 'usefulness' is the defining characteristic of true belief. For example, "the true is whatever proves itself good in the way of belief." At other times he seems content merely to stress the importance of this feature of true belief and to appeal to it as a motive for seeking true belief. This is the leading idea of the passage quoted above. Nowhere does he clearly formulate the claim that 'the truth pays.'

James' equation of true belief with useful belief was ridiculed by Russell and Moore. For example, Moore remarks that

> Is it not clear that we do actually sometimes have true ideas, at times when they are not useful but positively in the way? (Moore, 'William James' "Pragmatism",' 1908)

And Russell comments that false belief may be useful:

> It seems perfectly possible to suppose that the hypothesis that (other people) exist will always work, even if they do not in fact exist. It is plain... that it makes for happiness to believe that they exist.... But if I am troubled by solipsism, the discovery that a

Synthese **43** (1980) 369–379. 0039–7857/80/0433–0369 $01.10.

belief in the existence of others is 'true' in the pragmatists' sense is not enough to allay my sense of loneliness. (Russell, 'James' Conception of Truth,' 1908)

There certainly seems to be merit in these objections. It is easy to think of examples in which someone believes a truth which he would have been better off not to have believed, and conversely, situations in which a false belief results in beneficial consequences. But it is difficult to assess the force of these criticisms until James's equation of useful belief with true belief is made precise. In particular, some way of measuring the utility of a belief needs to be devised.

A number of philosophers have observed that there are obvious connections between various pragmatist doctrines and contemporary decision theory.[2] However, hardly anyone has attempted to formulate pragmatist views in decision-theoretic terms. An exception is Nicholas Rescher, who in *Methodological Pragmatism*[3] suggests a decision-theoretic formulation of the position that the truth of a proposition is correlated with the utility of believing it. He does not explicitly define 'the utility of believing S' but his examples suggest the decision-theoretic formulation that I will give. Rescher argues that the pragmatist's claim that there is a correlation between truth value and utility is fundamentally defective. After explicitly defining the utility of a belief I will show that while a simple-minded equation of the true with the useful does not work, decision theory contains the resources to provide a partial vindication of James' central insight: that the truth pays. In fact, we will see that a precise 'cash value' can be placed on the value of true information, at least relative to a given decision problem.

Bayesian decision theory provides a framework for representing and solving decision problems.[4] A decision problem, D, is specified by a set of actions, Δ, a set of mutually exclusive and exhaustive propositions, \$, a probability distribution on \$, and a binary function U which assigns to each pair $\langle d_i s_j \rangle \in \Delta \times \$$ the utility u_{ij} of the consequence that would result from choosing d_i when s_j is true. The utility of d_i is $\Sigma\, P(s_j) \times u_{ij}$. Those acts in Δ for which $u(d_i)$ is maximum are called the Bayes acts in Δ. Bayes Rule, which is a consequence of the axiomatic developments of utility and probability, requires a decision maker to choose a Bayes act. (We will assume that there is a Bayes act in the decision problems we are interested in although there are decision problems for which no Bayes act exists.) The probability and utility assignments are 'subjective' in the sense that they represent the beliefs and preferences of a particular decision maker; other decision makers may assign different probabilities and utilities. The utility of accepting proposition Q for decision problem D when s_k is the true member of \$, $U^D_{sk}(Q)$, is computed as follows: calculate the probability distribution $P(s_j/Q)$. Using this probability distribution calculate the Bayes act in the decision problem. Suppose d_i is the unique Bayes act in Δ. Then the utility that would be obtained from choosing d_i when s_k is true is u_{ik} and this is $U^D_{sk}(Q)$. If there is more than one Bayes act in Δ relative to $P(s_jQ)$, $U^D_{sk}(Q)$ is the average of the utilities that would be obtained from these acts when s_k is true. To simplify matters I will assume that the Bayes act in Δ is always unique. For example, in the following decision problem the utility of $s_1 \vee s_3$ when s_2 is true is 2.

I

	s_1	s_2	s_3	s_4
d_1	7	2	0	2
d_2	0	6	1	1
d_3	2	2	4	4
	.25	.25	.25	.25

(The numbers in the boxes are utilities; e.g. $u_{14} = 2$. The numbers at the bottom of the columns are the probabilities $P(s_i)$.)

An important feature of this account is that the utility of a belief is determined entirely by what I will call its *rational* consequences, and not at all by the utility of other consequences which may be caused by the belief. For example, suppose that A is considering whether or not to ask Professor X to write a letter of recommendation for him. Relative to this decision problem, true information concerning the kind of letter X would write obviously has high utility. If, in fact, X has a low opinion of A and so would write a damning letter, it is better for A to find this out before deciding to ask him to write the letter. If true, the belief that X has a low opinion of A is of high utility relative to this problem. However, this belief might also cause A considerable psychological stress so that A's overall state of utility is lower after discovering X's opinion of him. James seems to be aware that a belief may have psychological as well as rational consequences, but it is not clear as to whether the utility of a belief includes in his view psychological consequences. His discussion of the idea of the Absolute (p. 41) suggests that he considers the proposition that the Absolute exists to be true in so far as it brings comfort to those who believe it. In any case, he seems in this instance to include some psychological consequences, the comfort felt by one who believes in the Absolute, in assessing the utility of the belief. But for any proposition Q it is possible that believing Q may cause distress (or pleasure) to a person quite independently of its truth value. Believing Q might remind him of some horrible incident of his youth, thus upsetting him even though Q has nothing to do with this incident. So, if psychological as well as rational consequences were to be included in measuring the utility of a belief, there is no reason to expect there to be a connection between truth and utility. Our decision theoretic fomulation, since it takes into account only the rational consequences of belief, is an improvement over James' vague formulations.

Rescher formulates a version of the pragmatic theory of truth as follows: a statement S is true iff (1) its acceptance is itself of relatively high utility and (2) its acceptance is of greater utility than the acceptance of its denial. It is far from obvious that James' writings support this formulation. While James certainly stresses that true belief is more satisfactory than false belief he does not attempt to give necessary and sufficient conditions for truth as Rescher does. In fact, he remarks that he never intended that satisfactoriness of belief should be taken as a sufficient condition for truth (p. 272). In any case, since the measure of the utility of a proposition is relative to a decision problem, (1) and (2) are obviously inappropriate since they are not so relativised. James himself points out that a certain truth might be irrelevant and so worthless relative to a particular decision problem (p. 111). But the same truth will be valuable with respect to other decision problems. For these reasons the theory of truth embodied in (1) and (2) is not plausibly attributable to James.

Rescher has little difficulty refuting the contention that (1) and (2) are necessary and sufficient conditions for the truth of a statement. He shows the following: (i) there are decision problems for which the utility of a false proposition is greater than the utility of a true one. In particular the utility of a false proposition may be greater than the utility of its negation; (ii) there are decision problems in which a false proposition is maximally utile; and (iii) high utility does not have the logical properties of truth. For example, the conjunction of two propositions of high utility may have low utility, and some of the logical consequences of a true proposition of high utility may have low utility.[5]

These results not only demolish the account of truth based on (1) and (2) but they also seem to sever the kind of connection between truth and utility that James took to be the core of his pragmatist account. Responding to Russell, James wrote

> Good consequences are not proposed by us merely as a sure sign, mark, or criterion, by which truth's presence is habitually ascertained, tho they may indeed serve on occasion as such a sign; they are proposed rather as the lurking *motive* inside of every truth-claim... (p. 146)

But if the utility of a true proposition may be less than the utility of its negation it would seem that the attainment of high utility is not a reasonable motive for seeking true belief.

I agree with Rescher's conclusion that high utility is not a defining characteristic of truth. (But we saw that there is reason to doubt that James held this view). However, I will argue in Part II that there are some interesting connections between truth and utility which support James' contentions that the truth pays and that this provides a motive for seeking truth.

II

Consider this example discussed by Rescher:

> ... a man mistakenly believes that he has a certain disease and mistakenly believes a certain medication is called for by way of treatment. Unbeknown to himself, he actually has another, more serious malady against which this medicine is highly

effective. If he 'knew the truth' he would suffer badly, being altogether ignorant of any remedy for his actual condition.[6]

This example is intended to show that false belief may issue in better consequences than true belief. However, notice that the false proposition 'I have disease X and medication Y treats disease X' logically implies the true proposition 'medication Y treats whatever disease I have' and that this proposition has high utility. In fact, the false proposition has the high utility it does precisely because it implies the true proposition. This observation is completely general. So, for example, in decision problem I,[5] s_4 implies the true $s_2 \vee s_4$ and has the utility it does precisely because the utility of doing d_3 when s_2 is true is 2.

It is not difficult to see that if s is false and s_k is the true state of nature, then $U^D_{s_k}(s \vee s_k) \geqslant U^D_{s_k}(s)$. The proof is this: suppose s_k is true. Then $U^D_{s_k}(s) = u_{j_k}$ where j is the index of the Bayes act, d_j, in D computed relative to the probability distribution $P(s_i/s)$. $U^D_{s_k}(s \vee s_k)$ is calculated in the same manner except the probability distribution $P(s_j/s \vee s_k)$ is used. If d_j is the Bayes act in D relative to the distribution $P(s_i/s \vee s_k)$ then $U^D_{s_k}(s \vee s_k) = U^D_{s_k}(s)$. But if some other act, d_h. is the Bayes act then u_{h_k} must be greater than u_{j_k}. In this case $U^D_{s_k}(s \vee s_k) > U^D_{s_k}(s)$.

This result means that if a false proposition has high utility it does so only in virtue of entailing a true proposition of at least as high utility. In the decision–theoretic framework someone who assigns a probability of 1 to the false proposition s also assigns probability 1 to $s \vee s_k$. So, if acting on his false belief s earns him a high utility with respect to D, it is because he also has the true belief $s \vee s_k$. These considerations, it seems to me, go some distance toward re-establishing the connection between truth and utility. It provides an answer to the question of how a false theory can still be highly useful. For example, Ptolemaic astronomy (Rescher's example) is useful for predicting the positions of the planets, simply because it has true consequences concerning their positions. Relative to a particular decision problem, then, a certain false theory may be as good or almost as good as a true one. But the reason will be that the false theory has true consequences. Far from refuting James, these observations lend support to his claim that it is *the truth* which pays.

Even if false beliefs are beneficial only in virtue of the truth they contain, some truths, as Moore says, may be positively detrimental. For example, suppose that in I s_4 is the true state of nature. Acting without further information, a decision maker would choose d_3 and consequently obtain a utility of 4. But suppose that before acting he learns that $s_1 \vee s_4$ is true. $U^1_{s_4}(s_1 \vee s_4) = 2$: so learning this truth has the effect of reducing the utility obtained from 4 to 2. We might say that a person contemplating decision problem I would be better off if he were not to learn this particular truth. Does this observation confound the contention that the truth pays? First, notice that although a partial truth may be irrelevant or misleading relative to a decision problem, the whole truth, that is, the truth about which state of nature obtains, is always valuable. $U^D_{s_k}(s_k)$ is greater than or equal to $U^D_{s_k}(s)$ for every s. What the examples show is that, relative to a particular decision problem, a little learning can be a dangerous thing. Someone who knows the whole truth will be able to determine that $s_1 \vee s_4$ would be a misleading truth for a decision maker contemplating I to obtain. (More accurately, it would be misleading for him to obtain precisely the information $s_1 \vee s_4$. If he were to believe s_4, he would also believe $s_1 \vee s_4$, and his action would obtain a utility of 4). However, from the point of view of a decision maker confronting decision problem I, learning the truth value of $s_1 \vee s_4$ will appear to be valuable. Recall James' remark that good consequences are 'the lurking motive inside of every truth claim.' He seems to mean that the prospect of good consequences is the motive which underlies the search for truth. Decision theory provides a way of precisely formulating and demonstrating this.

Consider the following decision problem:

II

	s_1	s_2
d_1	10	0
d_2	0	8
	.25	.25

In this problem the expected utility of d_1, $u(d_1)$, is 5 while $u(d_2) = 4$. Suppose that before making his decision A could find out which of s_1 or s_2 is true. We can imagine an oracle that will infallibly supply him with information concerning the true state of nature. How valuable is this information to A? The problem can be analyzed as follows: the oracle will either report that s_1 is true or that s_2 is true. If the former, A will choose d_1 just as he would have had he not consulted the oracle. For this reason we will call the answer s_1 'ineffective'. In contrast, if the oracle asserts that s_2 is true, then A will choose d_2 and obtain a utility of 8 instead of the utility 0 which he would have obtained had he acted without consulting the oracle (since he would then have chosen d_1 which would have resulted in obtaining 0 utility since s_2 is true). So his net gain in utility is in this case $8 - 0 = 8$. Before consulting the oracle, A assigns a probability of .5 to the oracle saying 's_1 is true', and an equal probability to its saying 's_2 is true', since these are the probabilities he assigns to s_1 and s_2 and he knows that the oracle is infallible. But this means that A's expected gain in utility to be obtained from consulting the oracle is $.5 \times 8 = 4$. This is called the *value of perfect information* for this decision problem. More generally, the value of perfect information for a decision problem is given by the expression

$$\sum_{i}^{n} \mathrm{P}(s_i)(\max_{j}(u_{ji}) - u_{ki}).$$

In this expression u_{ki} is the utility that results from performing the act with the greatest expected utility when s_1 is true.

The expected utility of perfect information is never less than 0; and, as long as $u_{ij} \neq u_{xy}$, for some i, j, x, y, it is strictly positive. The expected value of perfect information is the maximum amount it would be rational for A to pay, in terms of utility, for true information concerning which state of nature obtains. Since this value is typically greater than 0, decision theory does justify the claim that the whole truth is valuable.

Suppose now that the oracle is not infallible, that it sometimes provides false information concerning the state of nature. For example, the probability that the oracle tells the truth concerning which of

s_1, s_2, holds is .8. Letting e_i stand for 'the oracle says s_i is true', this is expressed by $P(e_i/s_i) = .8$. Consulting the oracle is like performing an experiment with outcomes e_1, e_2 and error probabilities $P(e_2/s_1) = .2$ $P(e_1/s_2) = .2$. As might be expected, the value of the oracle's information this time – it is called the expected value of sample information (EVSI) – is less than the value of perfect information. The value of consulting this fallible oracle can be calculated as follows: A will obtain either e_1 or e_2 as an answer to his inquiry. If the first, he will find himself confronting decision problem $D(e_1)$; if the second, he will be facing $D(e_2)$.

$D(e_1)$

	s_1	s_2
d_1	10	0
d_2	0	8
	.8	.2

$D(e_2)$

	s_1	s_2
d_1	10	0
d_2	0	8
	.2	.8

Observe that the only differences between these decision problems and the original problem are the probabilities assigned to the states of nature. The probabilities in $D(e_1)$ and $D(e_2)$ are the conditional probabilities $P(s_i/e_1)$ and $P(s_i/e_2)$ which can be computed by using Bayes' theorem. The outcome e_1 is ineffective, since if A were to hear this news from the oracle, he would perform d_1 – the act he would have chosen without consulting the oracle. But if the oracle answers e_2, then A will choose d_2, which has an expected utility of 6.4. A would also recognize that had he chosen d_1, he would expect a utility of only 2. So if the oracle says e_2, then A experiences a gain in expected utility of $6.4 - 2 = 4.4$. Before consulting the oracle, A assigns a probability to e_2, which is calculated by

$$P(e_2) = P(s_1) \times P(e_2/s_1) + P(s_2) \times P(e_2/s_2) = .5.$$

So he can expect a gain in utility from consulting the oracle of $.5 \times 4.4 = 2.2$. This is the value of the sample information that can be expected to be obtained by performing the experiment of consulting the oracle.

In general, if $\mathscr{E}$ is an experiment with outcomes $e_1 \ldots . e_u$ the expected value of sample information (the expected value of the information that will be obtained from the experiment) relative to decision problem D is given by the expression:

$$\mathrm{EVSI}(\mathscr{E}) = \sum \mathrm{P}(e_i)(\mathrm{U}(d/e_i) - \mathrm{U}(d_k/e_i));$$

where $\mathrm{U}(d/e_i)$ is the expected utility of the act of maximum expected utility, computed by using the conditional probability distribution $\mathrm{P}(s_i/e_i)$. The summation is over the outcomes of $\mathscr{E}$. It is easy to see that EVPI $\geqslant$ EVSI $\geqslant 0$ for any problem D. Thus, the complete truth relative to D is most valuable. Since experiment $\mathscr{E}$ might lead to a false belief, the fact that EVSI $\geqslant$ EVPI supports the pragmatist position that it is the truth which is valuable, the less likely is it that the truth is to be obtained from $\mathscr{E}$ – the less valuable is D.

The disutility of false 'information' can be further seen as follows: suppose that A decides to choose d_i if and only if the oracle reports that s_i is true; and suppose also that the probability that the oracle tells the truth is p. The expected utility of acting on the oracle's report is then $1/2p \times 10 + 1/2p \times 8 = 9p$. If p is less than 5/9, then A would be better off acting without consulting the oracle, since the expected utility of acting without consultation is 5. In fact, if the oracle provides false 'information' with probability 1, then the utility of acting on the basis of this false 'information' is -5. Of course, it would be irrational to use the test knowing that the probability of misinformation is greater than 4/9.

The value of information which is known to be true but partial can be computed in the same way as EVSI. For example, the value to a decision maker confronting I of learning the truth value of $s_1 \vee s_4$ is equal to $\mathrm{P}(s_1 \vee s_4) \times (\mathrm{U}(d/s_1 \vee s_4) - 3) + (\mathrm{P}(s_3 \vee s_3)) \times (\mathrm{U}(d/s_2 \vee s_3) - 3) = .5 \times (4.5 - 3) + .5 \times (4 - 3) = 1.25$. Of course, if s_4 is the true state of nature, the decision maker will be worse off for having learned that $s_1 \vee s_4$ is true, as we have seen. This is explained by noting that if any of the other states of nature obtain then he will be better off acting after learning the truth value of $s_1 \vee s_4$. Since $\mathrm{P}(s_4) = .25$, from his point of view it is worth the risk resulting from acting on the information.

By casting James' views into a decision-theoretic framework, we have made precise and plausible his claim that

> ....the possession of true thoughts means everywhere the possession of valuable instruments of action; and that our duty to gain truth, so far from being a blanket command from out of the blue, or a 'stunt' self-imposed by out intellect, can account for itself by excellent practical reasons.

However, it is important to recognize that decision theory provides no support for the position that the pragmatic theory of truth is superior to correspondence, coherence, or other theories of truth. On the contrary, Rescher (following Russell and Moore) has convincingly shown that high utility cannot plausibly be construed necessary and sufficient for true belief. James himself, at times, seems to have taken his pragmatic theory of truth, not as a rival to these other theories, but as an attempt to show why truth is generally valuable. For example, he remarks that "the existence of the object, whenever the idea asserts it 'truly', is the only reason, in innumerable cases, why the idea works successfully, if it works at all" (p. 174). Here he seems to be assuming a correspondence account of truth and claiming that it is precisely because a truth corresponds to reality that it is useful. What I have done in this paper is use some simple results in decision theory to show that James' insight that the truth pays is fundamentally incorrect.

University of Southern Carolina

NOTES

* I would like to thank Ferdinand Schoeman, Michael Gardner, and Roger Rosenkrantz for helpful comments on earlier versions of this paper.

[1] William James, *Pragmatism* (Cambridge: Harvard University Press, 1978), p. 98. Page references are to this edition of *Pragmatism and the Meaning of Truth.*

[2] For example, Isaac Levi, *Gambling with Truth* (Cambridge, Mass.: M.I.T. Press, 2nd edition, 1967).

[3] N. Rescher, *Methodological Pragmatism* (Oxford: Blackwells, 1977), Ch. IV. Also N. Rescher and T. Vici, 'On the Truth-Relevancy of the Pragmatic Utility of Beliefs', *Review of Metaphysics* **28** (1975), 443–452.

[4] For an introduction to decision theory, see D. V. Lindley, *Making Decisions*, (New York: John Wiley, 1973).

[5] These results can all be illustrated by decision problem I.

(i) $U^{I}_{s_4}(s_3) > U^{I}_{s_4}(\sim s_3)$;

(ii) $U^{I}_{s_4}(s_3) = 4$;
(iii) $U^{I}_{s_2}(s_1 \vee s_2) > U^{I}_{s_2}(s_2)$.

[6] Rescher, *Methodological Pragmatism*, p. 56.

[7] For a discussion of the value of information, see Lindley, op. cit., or Roger Rosenkrantz, *Inference, Method, and Decision* (Dordrecht: Reidel, 1977).