Attitude Ascription's Affinity to Measurement

Mitchell S. Green

Abstract

The relation between two systems of attitude ascription that capture all the empirically significant aspects of an agent's thought and speech may be analogous to that between two systems of magnitude ascription that are equivalent relative to a transformation of scale. If so, just as an object's weighing eight pounds doesn't relate that object to the number 8 (for a different but equally good scale would use a different number), similarly an agent's believing that P needn't relate her to P (for a different but equally adequate interpretive scheme could use a different proposition). In either case the only reality picked out by any system of ascription is what is common to all equivalent rivals. By emphasizing some contrasts between decision theory and belief-desire psychology, it is argued here that if attitude ascription is appropriately analogous to measurement, then not only is being related to a proposition an artifact of the system of representation chosen, but so are belief and desire.

1 Measurement and its artifacts

Non-eliminativism about propositional attitudes does not, it has been said, require their construal as relations to propositions. Without attempting to resuscitate adverbial or sententialist accounts of such ascriptions, some recent authors have contended that an agent's believing that P no more requires her to stand in a relation to P than does an object's weighing five pounds require it to stand in a relation to the number 5. So long as the chosen measurement scale keeps track of the relevant empirical relations among objects (such as balancing on an equal arm balance), it does not matter which particular numbers are used as indices of those relations. Similarly, so long as a system of attitude ascription keeps track of such phenomena as an agent's preferences among outcomes or dispositions to assent, it does not matter which propositions are used to index

International Journal of Philosophical Studies Vol. 7 (3), 323–348
ISSN 0967–2559 print 1466–4542 online

them. Because the choice of a number to index a magnitude (proposition to index thoughts or behaviour) is within certain limits arbitrary, the urgency of the idea that in having a magnitude an object bears a relation to a number (in thinking or behaving thus and so the agent bears a relation to a proposition) is lessened. In spite of their relational surface grammar, sentences reporting attitudes need not, according to this line of thought, be given relational truth conditions. By challenging the question's presupposition, this conclusion may exonerate us from answering the question how standing in a relation to a proposition can be causally efficacious.[1]

Some definitions will ease formulation of the mooted analogy. Say that *U is a relational system* iff U is a finite sequence $<A, R_1, \ldots, R_n>$, where A is a (non-empty) set of elements (called the *domain* of U) and $R_1, \ldots, R_n$ are relations on A. Where $U = <A, R_1, \ldots, R_n>$ and $U' = <B, S_1, \ldots, S_n>$ are relational systems, *U′ represents U* iff U′ is a homomorphic image of U, that is, iff there is a function f from A into B such that for all $x_1, \ldots, x_k$ in A, $R_i(x_1, \ldots, x_k)$ implies $S_i(fx_1, \ldots, fx_k)$. U is a *numerical relational system* iff the domain of U is a set of real numbers, and U is a *formal relational system* iff its domain consists of abstract objects such as numbers, points, sets, etc. (Essential to measurement is the use of numerical relational systems to represent relational systems the elements of whose domain are given empirically.) Once a relational system has been given a formal representation, the question arises to what degree that representation is unique. Where a formal relational system $U' = <B, S_1, \ldots, S_n>$ represents relational system $U = <A, R_1, ..., R_n>$ by means of the mapping f, a transformation T from B onto itself is admissible with respect to f just in case $T[f(x)]$ also bears witness to the representation of U by U′. The set of all admissible transformations determines the scale type or the degree of uniqueness of the system of measurement chosen. (Among the better-known scale types definable on numerical relational systems are *nominal scales*, on which all one–one transformations are admissible, *ordinal scales*, on which all monotone transformations are admissible, *interval scales*, on which all positive linear transformations are admissible, *ratio scales*, which are uniquely determined except for an arbitrary unit of measurement, and *absolute scales*, on which no transformations are admissible (an example is counting).)[2]

It is natural to hold that where T is an admissible transformation of f, then if $f(x) \neq T[(f(x)]$ the particular entity onto which f maps x is to some extent an artifact of the system of representation chosen. This is perhaps why assigning the number 5 to an object as its weight need not be understood as placing it in a relation – for instance, the weight-in-pounds relation – to that number. If it were to be so understood, then by virtue of weighing-in-pounds what it does the object would stand in a relation to a great many other numbers as well. Such a view of measurement would also raise the question how an object's having the weight-in-pounds that it does, construed as its standing in a relation to an abstract entity, could be causally efficacious.[3] Rather, following Suppes and Zinnes, 1963 we may say that what is, or at least purports to be, objectively correct in any system of formal representation is what is common to all admissible transforms thereof. For instance, instead of holding that a non-absolute measurement scale relates objects to particular numbers, we may hold that any such scale ascribes features to objects by indexing these features with abstract entities that preserve the significant aspects of those objects and their interrelations.[4]

Those who contend that more than one system of intentional ascription can capture all the significant features of an agent's thoughts and speech acts may see an analogy here. Suppose, for instance, that only one of two empirically adequate systems of attitude and speech-act ascription imputes to an agent the belief that P, the other ascribing no belief with that content. One of these systems of ascription may nevertheless be related to the other in a manner analogous to that in which one measurement system is related to another that results from it by means of a scale transformation. If so, then since the two systems of intentional ascription differ over whether the agent in question believes that P, the ascription of that belief must, it will be suggested, be an artifact resulting from the choice of one system of ascription over the other. It is not *false* to say of the agent that she believes that P, any more than it is false that her body temperature is 98.6 degrees Fahrenheit. Rather, both ascriptions contain elements that are, strictly speaking, innocuous but able to tempt us into unwarranted ontological conclusions.

In helping us to discern the line between the factual and the artifactual, the study of admissible scale transformations and possible interpretive schemes helps to guard us against such temptation. In the temperature case we may clearly exhibit that line by characterizing the set of admissible transformations for any system of temperature measurement, and an attempt to delineate the distinction between the factual and artifactual in attitude ascription may also be of value. For if we are mindful that all that purports to be objectively correct in any system of attitude ascription is what is common among all empirically adequate rivals, we shall not be forced to infer that the agent stands in a relation to the proposition P just from the fact that one empirically adequate system of attitude ascription imputes to her an attitude with that content. At the very least that inference awaits the finding that all empirically adequate systems of ascription do so. Davidson has put the point this way:

> The same facts may be represented by quite different assignments of numbers. In the interpretation of speech, introducing such supposed entities as propositions to be meanings of sentences or objects of belief may mislead us into thinking the evidence justifies,

> or should justify, a kind of uniqueness that it does not. . . . Propositions being much vaguer than numbers, it is not clear to what extent they are overdesigned for their job.
>
> (Davidson, 1974: p. 147; see also his 1980b: p. 6, and 1989: pp. 9–11)

The analogy between attitude ascription and numerical representation (hereafter *the measurement analogy*) is unlikely to admit of quick validation or refutation. For one thing, unlike the case of, say, the measurement of mass, we have no concise and uncontroversial account of the relations on the empirical domain that belief/desire ascription is intended to preserve. This lacuna need not vitiate the measurement analogy – it was, after all, only in 1901 that Hölder explicitly characterized the empirical properties and relations that the measurement of mass is to capture – but it does hamper its assessment. Rather than presuming to fill this lacuna, in what follows I shall navigate a slightly roundabout route to the assessment of the analogy by showing how it admits of a natural development that has apparently gone unnoticed by its proponents. Mathematical psychology has long treated decision theory as sharing affinities with measurement in the physical sciences, and, as will emerge, certain forms of decision theory rival belief-desire psychology in their ability to provide abstract objects that index the interrelations among cognitive and conative states. As will be argued, however, these forms of decision theory may do this without ascribing beliefs and desires. By adverting to the relation between one such version of decision theory, known as prospect theory, and belief-desire psychology, it is argued here that if the measurement analogy is apt, then not only is being related to a proposition an artifact of the system of representation chosen, but so are the very characteristics of having beliefs and desires. Without establishing the artifactuality of propositional attitudes, which are ascribed by both decision-theoretic and belief-desire systems, this conclusion suggests that the measurement analogy may be used to show beliefs and desires to be as much artifacts of the representational system used as are the relations they are often taken to stand in to propositions. Whether such a development of the measurement analogy reduces it to absurdity or, on the other hand, vindicates it will not be settled here. Rather, whichever of these possibilities obtains, it may be hoped that the present development of the analogy will contribute to determining its aptness.

2 Attitude ascription as analogous to measurement

In an early discussion of the analogy between attitude ascription and measurement, Field (1981) develops a suggestion of David Lewis' that the domain of the empirical relational system that attitude ascription may be seen as representing comprises internal states having sentence-like properties. One virtue of this approach is that it suggests an account of what properties and relations in the empirical domain are to be preserved in the process of representation. On this approach, if, for instance, internal state S, in virtue of its syntactic properties, syntactically implies internal state S′, in virtue of *its* syntactic properties, then any formal system representing this relation by means of a mapping f must be one in which $f(S)$ logically implies $f(S')$. If the domain of the empirical relational system is construed as a set of sentence-like entities, then, as Field suggests, we shall have some guidance in determining what are the relevant properties in the empirical relational system requiring to be preserved in representation.

Field takes these considerations to support a view of the internal states to be represented by a formal relational system as having sentence-like properties, but other proponents of the measurement analogy have demurred. Stalnaker (1984: Chapter 1), for instance, argues that the relations in the domain of the empirical relational system that are to be preserved in any formal representation are causal and functional characteristics of those states. One of those states, which would be represented in the formal relational system as the desire that P, disposes its possessor to act in such a way that she brings it about that P in a world in which her beliefs, whatever they might happen to be, are true. Another of these states, which would be represented in the formal relational system as the belief that P, disposes its possessor to act in such a way that she satisfies her desires, whatever they might happen to be, in a world in which P (as well as her other beliefs) are true.[5] It does not seem, however, that such functional characterizations of these states require imputing to them linguistic structure, and Stalnaker argues at length (1984: Chapter 2) that such an imputation is not compulsory.

Matthews, 1994 develops a similar line, contending that the domain of the empirical relational system that attitude ascription may be seen as representing comprises those internal states of organisms that are causes of their optimizing behaviour. On Matthews' account, the principle guiding our mapping of those states onto the representational space used in attitude ascription is that the properties preserved in the mapping, 'are those causally efficacious properties of internal states in virtue of which possessors of these states usually succeed in behaving optimally in their environment, i.e., in a way that satisfies their needs' (pp. 139–40). He infers from this that to a first approximation an internal state that is shaped by the environment is represented by the system of belief-desire psychology as belief, whereas one that shapes the environment is represented by that system as desire.[6] We ascribe contents to those states as a way of indexing the distal environmental situations that shape and are shaped by them; characterizing an internal state of an organism as a belief that P, then, tells us that it is one that is typically brought about by the state of affairs

P, while characterizing an internal state of an organism as a desire that Q tells us that it is a state that typically brings it about that Q. The representational system available from belief-desire psychology does not itself incorporate any postulate linking attitudes and action, but it may be used in conjunction with such logically independent postulates as that agents strive to fulfil their desires given that the world is as they take it to be, to allow users of that scheme to rationalize behavior.

We need not here take a stand on the dispute between Field, on the one hand, and Stalnaker and Matthews, on the other, concerning the amount of sentence-like structure to be imputed to items in the domain of the empirical relational system that is represented in attitude ascription. So too we may remain neutral on the question of the exact structure of the formal relational system that will represent attitudes. All proponents of the measurement analogy agree that the formal relational system must contain propositions, but they will differ over how much, if any, internal structure must be imputed to propositions. Just to fix ideas and to have a specific example with which to work, however, let us observe that Matthews (1994: pp. 135–6), following Richard's (1990) synthesis of Russellian and sententialist conceptions of attitudinal content, contends that the space onto which items in the empirical domain of belief-desire psychology are mapped comprises ordered pairs $<a_i, <s_j, r_k>>$, where a_i is an attitude type, s_j a sentence type and r_k a Russellian proposition (itself an ordered n-tuple consisting of n–1 individuals and a property (or relation)) capable of being expressed, relative to a context of utterance, by s_j. This approach allows the items used to index attitude states to individuate them at least as finely as do sentences, while still representing those states as contentful. Observing that not all aspects of a representation space can be read back into the empirical domain that it represents, Matthews (pp. 138–9) leaves it open whether a function mapping internal states of organisms into a space comprising propositions justifies the conclusion that all inferential relations among propositions have an image in relations of internal states to one another. Consequently, such failures in agents as departures from logical consistency in beliefs and from deductive closure in beliefs do not provide an immediate objection to the empirical adequacy of belief-desire psychology as here construed.

Matthews does not state, much less prove, a representation theorem for belief-desire psychology, perhaps because it is insufficiently clear which relations among internal states this form of psychology is intended to capture.[7] On the question of admissible transformations, Matthews argues that whether two representations are of the same propositional attitude depends on whether, in ascribing an attitude, our explanatory interests require us to be faithful to the ascribee's words or point of view; he concludes that an account of the degree of uniqueness of the system of representation employed by belief-desire psychology must invoke pragmatic considerations. However, this view is not required for adherence to the measurement analogy, even after the formal relational system's domain has been characterized as comprising propositions as conceived by Richard. In an attempt to characterize the space of admissible transformations one could instead advert to the indeterminacy confronting, say, Davidson's radical interpreter. From such a perspective an account of the degree of uniqueness of attitude ascription need not invoke pragmatic considerations. Rather, on such an approach only a criterion of empirical adequacy need be invoked in determining whether one system of attitude ascription is to be construed as an admissible transform of another.

Consistently with other authors who have taken up the measurement analogy, Matthews conjectures that attitude type will always be preserved by admissible transformations: even if one system of attitude ascription imputes to Smith the belief that P while an admissible transform of that system ascribes to him the belief that Q, any admissible transform of either system will construe the internal state in question as a belief. Matthews is right to enter this as a conjecture rather than assume it as self-evident. After all, it is not self-evident that the only causally efficacious properties of internal states in virtue of which possessors of these states usually succeed in behaving optimally in their environment can be the properties of being beliefs or desires. Among other candidate properties are those in virtue of which agents may be seen as maximizers of utility.[8] In what follows, Matthews' conjecture that attitude type will be preserved by admissible transformations will be challenged in light of scrutiny of some conceptual discontinuities between belief-desire psychology and decision theory.

3 Subjective expected utility theory and belief-desire psychology

It is sometimes said that decision theory is belief-desire psychology made precise by means of being made quantitative,[9] and while this view is not wholly incorrect, it can occlude some dramatic differences between the two kinds of theory. Because of its relative familiarity I shall focus on the relation between subjective expected utility theory and belief-desire psychology; another more empirically adequate form of decision theory will be mentioned presently. Subjective expected utility theory begins with the attribution to an agent of a set of available acts, a probability function P defined over a set X comprising mutually exclusive and jointly exhaustive states of the world possible by the agent's lights,[10] and a utility function U defined over the consequences of those acts depending upon which state of the world may happen to obtain.[11] These notions are the elements of subjective expected utility theory's mapping of internal states of an organism that are causal antecedents of its optimizing behaviour onto points in a representational space; here the points may be construed as propositions paired with real numbers in the unit interval (for degree of belief), or by cardinal numbers (for utility assignments to outcomes).

As with the mapping provided by belief-desire psychology, these notions do not by themselves provide a general means for the prediction and explanation of action, but that means is available from an independent postulate. Let 'a(x)' denote the consequence that will be obtained if act a is chosen and state $x \in X$ obtains; then so long as X is countable we define the *subjective expected utility* of an act a thus:

$$SEU(a) = \sum_{x \varepsilon X} P(x)U(a(x))$$

Subjective expected utility theory has it that an agent performs that act, if there is one, whose subjective expected utility is at least as great as that of any other act available; where there is more than one such act, the theory predicts only that one of them is chosen. This principle guides subjective expected utility theory's mapping from internal states onto the formal relational system just described, and is the basis of well-known representation and uniqueness theorems.

Subjective expected utility theory does not, in addition to the apparatus described above, also require the attribution of beliefs and desires; nor are they implicit in what has been described. There seems little reason to deny that a degree of belief is a propositional attitude, since it is an intentional state having propositional content. The same may be said for degrees of desirability. Nevertheless, in spite of the superficial similarity between such locutions as 'belief' and 'degree of belief', a degree of belief is no more a kind of belief than is an alleged criminal a kind of criminal; likewise for desire and degree of desirability. For one thing, while some of our beliefs are held more firmly than others – I believe that it will rain tomorrow less strongly than I believe that I will be conscious on that day – yet it is tendentious to express this by saying that belief is a matter of degree, for this suggests that all there is to believing that P is having some degree of belief that P. My degree of belief that I shall throw heads with the fair coin in my hand is .5, but I do not believe that I shall throw heads. (Nor need I believe that I won't; I might just be agnostic.) Hence a degree of belief is not *as such* a kind of belief.[12]

Nor, as the following case seems to show, could a degree of belief r in proposition P be identified with the belief that the objective chances of P are r. A box contains two coins, one that an agent is sure is biased so as to have a 2 in 3 chance of landing heads; the other that the agent is sure is biased so as to have a 2 in 3 chance of landing tails. She randomly picks a coin without seeing which one it is and flips. Her degree of belief that the coin will come up heads is .5 even though she does not believe that the objective odds of that coin's landing heads is ½.[13]

As has often been remarked, it also seems inadequate to define belief as degree of belief greater than or equal to r, for some r sufficiently close to but less than 1. For any r that might be proffered on behalf of such a definition, imagine a lottery I know to be fair and having sufficiently many tickets so that for each ticket t, my degree of belief that t will be a loser is greater than r. It follows that on the suggested definition, I also believe of each ticket that it will lose. This would seem to commit me to the conclusion that no ticket will be a winner, contradicting my knowledge that the lottery is fair and that therefore some ticket will win.[14]

Hunter, 1996 defines belief as sufficiently high subjective probability, but with the innovation that what counts as sufficiently high is vague. Belief attributions are according to this view only true to a certain degree, where being true to degree 0 is identified with being definitely false and being true to degree 1 is identified with being definitely true. On Hunter's definition,

> the degree of truth of 'x believes P' is $(2Pr(P) - 1)^n$, where Pr() is x's subjective probability function and n is an adjustable parameter greater than or equal to one.
>
> (1996, p. 87)

The context in which this definition is formulated suggests that the author means to stipulate that where $Pr(P) < .5$, $(2Pr(P) - 1)^n = 0$, for all n, and if this were not so the formula would yield bizarre results for values of P less than .5 when n is even. (Letting $n = 2$, we get the result that the degree of truth of 'x believes P' for $Pr(P) = .2$ is greater than the degree of truth of 'x believes P' for $Pr(P) = .75$.) It follows from this stipulation that no one believes any proposition to which she assigns degree of belief less than .5. But of a six-sided die I know to be fair, I might well believe that it will come up three on the next throw. Similarly, having just authored a large historical treatise I might recognize that, human fallibility being what it is, the proposition that is the conjunction of all propositions asserted in that book is very unlikely to be true. Yet I believe each of the propositions asserted in the book, and, recognizing this fact, I believe their conjunction.[15] Hunter's proposal, however, is inconsistent with the possibility of believing a proposition in which one has a degree of belief less than .5, and for this reason is unacceptable.

We can imagine an agent having nothing but relatively low degrees of belief for all the atomic propositions that can be formed out of her outcome space. That agent might fail to believe any such proposition (though, as we have just seen, this need not be so). Nor need the high degree of belief that is required on pain of probabilistic inconsistency for certain propositions formed by disjoining many of these atomic propositions in the outcome space issue in belief. For, as we have just observed, even a person's having high degree of belief in a proposition does not imply that she believes that proposition.

Remarks analogous to the foregoing may be made concerning the relation between desire and preference. Like beliefs, some desires are stronger or are held more firmly than others. Further, if it is possible strictly to order my preferences so that for every consequence a and b I either prefer a to b or *vice versa* or am indifferent between them, then one may assign numbers to all possible consequences so that utilities for them are given on an interval scale, invariant under linear transformations. (This is just to say that under such conditions my preferences can be represented by an interval scale.) Yet from the fact that a consequence has a ranking in a preference ordering it does not follow that I desire it. Nor does this follow if the consequence has a degree of desirability at least as high as any possible consequence of any available act. In a situation in which my only options are the loss of one hand and the loss of two, I prefer the former consequence over the latter without desiring to lose either hand. Likewise, on Davidson's original use of 'pro-attitude' it is required that one who bears such an attitude toward a consequence sees something desirable, just, or dutiful about that consequence. But even the loss of one hand in the just-imagined situation need not be construed in any of these ways.[16]

Some philosophers have come to use 'desire' in such a way that from the fact that an agent performs an act intentionally it follows analytically that she desired to perform that act.[17] On this usage a proponent of a decision-theoretic approach to behaviour who also takes the agents to which she applies her theory to perform intentional actions is committed to the view that those agents have desires. On the other hand the fact that an agent has a desire, in this wide sense, towards a certain consequence does not imply the existence of desires in the narrower, more colloquial sense of that term; nor indeed does it imply the existence of a conative state, internal or otherwise, that was causally efficacious in her behaviour.[18] As a result the decision theorist may accept philosophers' stipulation of this wider sense of 'desire' without undertaking any new commitments (other than what follow from this verbal one).

We may put these points more abstractly with the help of Stalnaker's (1984: pp. 79–81) notion of an acceptance state: a state *S is an acceptance state relative to proposition P* iff S is a state of an agent and S is correct iff P is true. While it is not clear that all cognitive states posited in common-sense psychology are acceptance states, a wide variety of them are. (Examples include belief, supposition, conjecture, and presumption.) Yet just in virtue of having some degree of belief in a proposition P an agent need not be in an acceptance state with respect to P. For in having some degree of belief in P the agent's state is not thereby correct if P turns out to be true. Normative notions can, to be sure, be used to assess an agent's degrees of belief. Many have held that certain patterns of degree of belief (both synchronic and diachronic) are criticizable if they leave the agent open to a book being made against her. Van Fraassen (1989: pp. 157–9) suggests that a forecaster's degree of belief r in proposition P is open to criticism if P fails to occur in r per cent of cases relevantly similar to those about which she makes that forecast. Yet while a degree of belief in a proposition may admit of a kind of vindication, it is one unlike what is due to the fact that belief aims at the truth.[19]

Now desire, which appears incapable of being correct, does not seem to be an acceptance state. However, a desire that P be the case is a state that has conditions of satisfaction, and in particular is a state that is satisfied just in case P holds. By contrast, a state of an agent can be construed as attaching a degree of desirability to a proposition P without being in a state that is satisfied just in case P holds. I attach some degree of desirability to the loss of one hand, but I am not on that account in a state that is satisfied just in case I lose one hand.

As we have seen, subjective expected utility theory involves ascription of more than degrees of belief and degrees of desirability. In order to make sense of an agent maximizing subjective expected utility we must also postulate (a) a set of possible acts open to the agent, (b) a set of mutually exclusive and jointly exhaustive hypotheses describing possible states of the world, and (c) for each act and each hypothesis, a description of what outcome that act will have if that hypothesis is true. (b) amounts to postulation of a space comprising all the outcomes the agent in question takes to be possible. In addition we must ascribe to an agent attitudes about what options are available to her, as well as views about what the outcomes of her acts will be depending what state the world is in. The latter are conditional probabilities, which with the help of Jeffrey's conceptualization thereof fit squarely within the probabilist framework. Similarly it should be clear that for the postulation of the set of possible acts open to an agent we need only require, for each such act, that the agent have some degree of belief that that act is an option for her.[20]

This leaves the outcome space. One might suggest that this can only be understood as comprising all the outcomes that the agent in question believes to be possible, and likewise for any pair of atomic elements in this space the proponent of subjective expected utility theory may seem committed to construing the agent as believing that they are incompatible with one another. The suggestions should be resisted. Concerning the first, the existence of an atom P in an outcome space is compatible with the agent's not believing that P is possible. For instance, given the right odds an agent could rationally wager in favour of a mathematical proposition P that she does not believe. To make sense of her behaviour the decision theorist would construe her outcome space as containing P. If, however, she does not believe that P, then she will presumably also not believe that P is possible.[21] In reply to the second suggestion, that a pair of atoms can exist in an agent's outcome space only if that agent believes them to be incompatible, it suffices to point out that we may capture an agent's views about the logical incompatibility of propositions P and Q by construing her outcome space as one in which the greatest lower bound of P and Q is the bottom element of the algebra.

4 A decision-theoretic alternative to subjective expected utility

We have seen that as a system for representing internal states by means of propositional attitudes a decision-theoretic approach need not invoke belief and desire. This does not imply that an agent's propositional attitudes can be rendered with no invocation of such folk-psychological concepts, but this stronger conclusion can be made plausible. First of all, the apparatus of degrees of belief and degrees of desirability as applied to states of the world and outcomes of actions, respectively, may be thought to be indirectly disconfirmed by the failure of agents to maximize subjective expected utility. Given a choice between a sure gain of $400 and a .5 probability of winning $1,000, many subjects choose the former even though its subjective expected utility is less than that of the latter.[22] This and many other well-documented, apparent departures from subjective expected utility maximization have led psychologists of decision in recent years to reject that organon as a descriptively adequate theory. On an holistic conception of theory confirmation the apparent descriptive inadequacy of subjective expected utility theory may be thought to undercut the apparatus of degrees of belief and degrees of desirability.

Recent years have, however, also seen the development of more descriptively adequate models of decision than subjective expected utility, one salient example being prospect theory.[23] On this approach the carriers of value, i.e., outcomes, are gains and losses rather than final assets, with the value function defined over these outcomes being (generally) concave for gains, convex for losses, and steeper for losses than for gains. The concavity for gains in the value function is designed to reflect subjects' tendencies to be risk averse for gains (preferring, e.g., a sure thing of $400 over a 50/50 chance of $1,000 or nothing). Convexity for losses is designed to reflect subjects' tendencies to be risk seeking for losses (preferring, e.g., a 50/50 chance of losing $1,000 or nothing to a sure loss of $400). Finally, the function's being steeper for losses than for gains is designed to reflect subjects' tendency to be loss averse (preferring the status quo to an uncertain prospect with the same expected utility).

In utility calculations these outcome values are multiplied not by additive probabilities but by so-called decision weights, which measure the impact of the personal probability of an event on the desirability of a prospect. The weighting function is designed to reflect such widespread tendencies as the exaggeration of small probabilities and underestimation of large probabilities as evinced in our decision making. According to prospect theory, the value of a given course of action, or prospect, is given by the sum of the products of the *values* of the various possible outcomes of that prospect with their *decision weights*. As in our brief exposition of subjective expected utility theory, we let 'a(x)' denote the consequence that will be obtained if act a is chosen and state x $\in$ X obtains, and we say that the value of prospect a is:

$$V(a) = \sum_{x \varepsilon X} w(P(x))v(a(x))$$

where P(x) is the agent's personal probability for an outcome x, w is the probability-weighting function, and v is the value function.[24]

In spite of its departure from subjective expected utility theory, prospect theory's refinement of the apparatus of degrees of belief and degrees of desirability as applied to states of the world and outcomes of actions, respectively, introduces no new primitives employing notions such as belief or desire; nor are these notions to be heard in the ensemble of concepts and techniques that it employs. Because of its basic distinction between gains and losses relative to the agent's status quo, prospect theory may presuppose a *de se* aspect of attitudinal content. It can do so, however, without presupposing beliefs, since degrees of belief can also be had towards contents that are essentially indexical. Again, it is tempting to take prospect theory's construal of certain outcomes as gains as a commitment to those outcomes being ones that the agent desires. (For how, it might be asked, could one fail to desire an outcome that is by one's own lights an overall gain?) The temptation is to be resisted, however, for whether a gain is desired can depend upon what one's relevant options are and her status quo. All those options might be horrendous by the agent's lights no matter what state the world happens to be in, particularly if she is in an already undesirable situation. In such a case it need not be that all, or even any, of the outcomes that are gains relative to the agent's status quo are ones that the agent desires.

5 Must decision theory be supplemented with belief-desire psychology?

Many behavioural decision theorists have found prospect theory to fare well in accounting for behaviour that seems not to maximize subjective expected utility.[25] It would thus appear premature to conclude that the apparatus of degrees of belief and degrees of desirability is implicated in the descriptive failure of subjective expected utility theory.[26] It would also be erroneous to view the belief-desire approach as an indispensable part of any comprehensive system of propositional-attitude ascription, including the decision theoretic. Some have, indeed, contended that the decision-theoretic apparatus cannot stand on its own as a way of rationalizing (and perhaps also predicting) choice behaviour, and that it requires supplementation with the apparatus of belief and desire. Thus Foley, discussing the point of view of an agent using decision theory to decide how to act, argues that when we make decisions in light of degrees of belief and desirability we do so, 'only against the backdrop of black-and-white assumptions – i.e., a backdrop of belief' (1992: p. 122; see also 1993: p. 171). One reason Foley gives for this is that when we make decisions we must first set up the parameters of the decision situation without qualification; we must have beliefs concerning the available acts, states, and outcomes in terms of which the decision problem is framed. In the last section we observed that some of these parameters, such as the acts the agent takes to be available, can be construed from the decision-theoretic point of view in terms of high degree of belief. Foley is aware of this possibility, but argues that we could not feasibly take all these background degrees of belief into account before deciding how to act. Instead, 'We are content with our black-and-white beliefs about these matters' (1992: p. 122; see also 1993: p. 172). Foley evidently intends not just that we are content with such beliefs but that we *must* rely upon them in deciding how to act.

It is indeed true that in some cases it is not feasible for an agent to calculate whether she has correctly framed a decision problem.[27] After all, such a calculation is an undertaking requiring time and effort, and might be performed in lieu of some more fruitful alternative. However, it is not clear why it should follow, from the unexceptionable premise that in some cases it is not feasible for an agent to calculate whether she has correctly framed a decision problem, that in such a case the agent must believe that she has framed her decision problem correctly. Imagine a situation in which it is not feasible for an agent to calculate whether she has correctly framed a decision problem, and in which she is agnostic about whether she has framed it correctly while having a high degree of belief (.9, say) that she has done so. The agent's being in such a situation need not *undermine* the formulation that she does have of her decision problem, and so need not prevent her using that formulation in determining how to act. Nor need a situation such as this prevent her from acting rationally (where rational action is defined either in terms of maximization of subjective expected utility or in terms of the machinery of prospect theory). If the agent were to have evidence undermining her high degree of belief that she has framed her decision problem correctly, that would tend to count against the rationality of her using that formulation of the decision problem to guide her action. Such a case, however, would seem precisely to be one in which it is rational to calculate whether her formulation of the decision problem is correct. Forming a belief that her initial formulation of the decision problem is correct might also help to allay her concerns, but nothing that Foley has established shows that the formation of such a belief is required either for action or for rational action.[28]

It does not seem, then, that the imputation to an agent of a formulation of a decision problem requires the further imputation of beliefs. Might the process of interpreting agents in the course of making sense of their behaviour do so? Davidson (1974) has argued that a 'radical' theorist of decision, who is like the better-known radical interpreter but who undertakes to build upon an austere basis a theory of utilities and subjective probabilities for an actor (rather than beliefs and desires), can only do so by communicating with that agent. If this is so, however, then the radical decision theorist must also be a radical interpreter. However, Davidson has also argued that we cannot tell what a speaker's words mean without knowing what she believes, and it would appear to follow that a radical decision theorist must also ascribe beliefs to the agent for whom she is constructing a theory of utility and subjective probability.

The claim in this argument that may be challenged is that we cannot tell what a speaker's words mean without knowing what she believes. It would seem to be enough for us to know what her words mean if we can tell what degrees of belief she has. More exactly, although Davidson (1974: p. 144) has proposed that the evidential base for a theory of radical interpretation is what sentences the speaker holds true at which times, he has also shown it to be equally workable to let the evidential base be what sentences the speaker holds true to at least degree r, where r is a fixed value on the unit scale well above .5. If, then, we can by triangulation solve for the meaning of the sentence, we shall be able to discern the agent's degree of belief in that sentence.[29] Belief need not enter the picture.

It might also be contended that a decision-theoretic approach to behaviour cannot make sense of the fact that we often, and with no apparent lapse of rationality, commit ourselves categorically on issues by making assertions that are not, or at least are not explicitly, qualified by probabilistic assessments. As Foley remarks, we often want others, 'whether they be scientists, teachers, butchers, journalists, plumbers or simply our friends, to act as jurors for us, delivering their black-and-white judgments about the facts as best they can' (1992: pp. 122–3). This is in part because of the heuristic value in delivering such 'black-and-white judgments' as compared with attempting to express a fine-grained degree of confidence for every proposition we utter without categorically asserting.

Theorists of decision need not deny that agents perform such speech acts (just as they may agree that agents make promises and sign contracts), and they may treat assertions as species of action, performed, when they are, in light of the appropriate utility calculation. Like contracts and promises, assertions engender commitments: one who asserts P is committed to giving reasons for P if challenged, and so long as no such challenge has been raised and gone unmet, is entitled to use P as fodder for further inference, i.e., as a basis for further assertions.[30] Acquiring the inferential status accruing to an accepted assertion evidently has value, and the prospect of gaining this status may in some cases justify the effort of making an assertion as well as the risk of being mistaken if that assertion turns out to be untenable. We may thus agree with Foley's contention that as interlocutors we often want categorical assertion rather than expressions of degree of belief, without relinquishing the contention that decision theory need make no crucial use of belief.

Without invoking the concept of belief the decision theorist can also make sense of the importance we attach to the distinction between assertions that are sincere and those that are not. Whereas on our ordinary conception, an assertion is sincere just in case the speaker believes what she asserts, from a decision-theoretic perspective the question of an assertion's sincerity divides into two: one who avows a degree of belief or desire is sincere just in case she actually has that degree of belief or desire; on the other hand for those assertions that are not also avowals of degrees of belief or desire sincerity is only a matter of degree, and varies according to the tenacity with which the speaker is willing to defend the asserted content, and the variety of contexts in which she is willing to act upon it. On this approach the homeowner who, in spite of asserting that her house will not flood during the impending summer, purchases flood insurance to protect herself against possible loss would count as less sincere than if she had spent that sum adorning her flowerbed. If, however, we are mindful that actions speak louder than words (even those said only in the heart), this need not come as an astonishing result.[31,32]

6 The artifactuality of belief and desire

In arguing that decision theory provides a means for representing agents' intentional states that does not presuppose the concepts of belief and desire, we have in effect characterized an enlargement of the domain of the relational system conjectured by proponents of the measurement analogy to represent certain of our internal states. Rather than containing only ordered pairs of propositions and indicators of belief or desire, we now construe that domain as comprising in addition ordered pairs of propositions and real numbers in the unit interval (for degrees of belief or, more precisely, decision weights) and ordered pairs of propositions and cardinal numbers (for degrees of desirability or, more precisely, increments of gain or loss). This enlargement of the domain of the relational system suggests the possibility of a larger space of admissible transformations than permitted by a measurement analogy restricting itself to belief and desire, with a concomitant thinning of what is in common among all points in that space.

To justify such an enlargement of the relational system's domain consider the example of two Martians, Brown and Jones, attempting so to map the internal states of an earthling, Ella, into a formal relational system as to characterize her as usually behaving optimally in her environment.[33] Suppose that Brown represents her with the tools of a well-confirmed behavioural decision theory such as prospect theory while Jones does so from the belief-desire perspective. A proponent of the measurement analogy, recall, takes it as a plausible conjecture that with world enough and time Jones could construct more than one system of belief-desire ascription characterizing Ella's propositional attitudes. This will motivate the conclusion that bearing the belief or desire relation to any particular proposition is likely to be an artifact of the system of representation chosen, thereby undermining a relational construal of attitudes. Since, however, by hypothesis Jones' aim was so to map Ella's internal states into a formal relational system as to characterize her as usually behaving optimally in her environment, an expansion of that relational system as proffered by Brown can make Jones no less able to achieve that aim, and will provide him with more ways of achieving it. Now some of Jones' systems of attitude ascription will employ only beliefs and desires; others available to him will employ in addition decision weights and increments of gain and loss; yet others will, after the fashion of Brown, employ only decision weights and increments of gain and loss. There may be no succinctly statable rule taking Jones from one system of ascription to an admissible transform thereof, but such an absence is no novelty for ascribers of belief and desire who contemplate indeterminacy.[34]

As we have argued, systems of the third sort (employing only decision weights and increments of gain and loss) do not ascribe beliefs or desires. Consequently, the proponent of the measurement analogy is in a position to hold that since only certain systems of propositional attitude ascription available to Jones ascribe beliefs and desires to Ella, these states are themselves to some extent artifactual.[35] Prising apart fact from artifact in attitude ascription not only calls into question a conception of propositional attitudes as relations to propositions, but also casts doubt on whether propositional attitudes must be construed in such terms as belief and desire. For while, as we have observed, degrees of belief and degrees of desirability are propositional attitudes, the former are not species of belief and the latter are not species of desire. The proponent of the measurement analogy may, accordingly, maintain a non-eliminitavism about propositional attitudes without postulating beliefs and desires.

Acceptance of the original measurement analogy does not, of course, imply commitment to its strengthening as here envisioned. Like the sapling to the seed, that strengthening is nevertheless a natural further development of the original analogy, which will therefore be supported or made less attractive depending upon the plausibility of the result. Which of these possibilities is realized cannot profitably be decided here. However, on behalf of the former possibility it may be noted that just as the measurement analogy in its original form offers, not so much an answer to the question how standing in a relation to a proposition can be causally efficacious, but rather a rejection thereof as containing a false presupposition, so too the strengthened analogy puts in a new light a traditional picture of ourselves, *qua* inquirers, as aiming at the truth and, *qua* agents, as aiming at satisfaction of our desires. On the approach of the strengthened measurement analogy it is not false to say of an agent that she believes that P or desires that Q, justifiable only as a bit of dramatic idiom or on purely instrumental grounds. Rather, like their contents belief and desire are to be seen as overdesigned for the job they perform in characterizing our propositional attitudes. Ascription of propositional attitudes no more requires belief and desire than does non-fundamental temperature measurement require use of the Celsius system. Consequently, from the perspective of the strengthened measurement analogy an insistence that our characterization as inquirers and agents requires a conception of ourselves as aiming at truth and satisfaction carries with it all the numerological virtues of an insistence that water on the boil is twice as hot as when it is 50 degrees Celsius. A therapeutic response to either way of thinking would exhibit the contingency of our practices.

University of Virginia, USA

Notes

1 The analogy between attitude ascription and measurement has been defended by Churchland, 1979, 1989; Leeds, 1979; Field, 1981; Stalnaker, 1984; Dennett, 1987 (Chapter 5), 1991; Davidson, 1974, 1989; Swoyer, 1987; and Matthews, 1994. Churchland, 1979 advocates an adverbial analysis of attitude ascriptions, but as Davidson 1989 remarks, this untenable position is logically independent of the measurement analogy.

Earlier versions of this paper were presented at the 1997 meeting of the Central Division of the American Philosophical Association, the University of Pittsburgh's Center for Philosophy of Science, the Central States Philosophical Association, and the University of Virginia Cognitive Studies Group. I am grateful to Mark Kaplan, John Barker, Fred Kroon, Michael Kubovy, Ram Neta, and Henry Jackman for their comments on those occasions. I would also like to thank the editor of this journal as well as two anonymous referees for their comments on an earlier draft.

2 For further development of these concepts see Suppes and Zinnes, 1963.

3 The point is well made by Crane, 1990: pp. 225–6.

4 Stalnaker writes, 'What is it about such physical properties as having a certain height or weight that makes it correct to represent them as relations between the thing to which the property is ascribed and a number? The reason we can understand such properties – physical quantities – in this way is that they belong to families of properties which have a structure in common with the real numbers. Because the family of properties which are *weights* of

physical objects has this structure, we can (given a unit, fixed by a standard object) use a number to pick a particular one of the properties out of the family' (1984: p. 9).

5 Stalnaker recognizes that while this account might serve to determine that a given internal state has the attitudinal type of belief, it is not adequate to determine that state's content. What is needed is not just a functional account of the connection of internal states with behaviour, but also an account of their production by states of the environment (Stalnaker, 1984: pp. 18–19). As will emerge presently, Matthews' development of the measurement analogy respects these constraints.

6 For further development of this approach see Dretske, 1988, particularly Chapters 3–6. We note also that on the present formulation it is not the case that the items in the empirical domain are assumed in the process of formal representation to be beliefs or desires. It might turn out that any adequate representation of those items by means of a formal relational system will map them onto entities having belief-like or desire-like properties, but that conclusion would be the upshot of a general account of successful representation and not an initial desideratum for any adequate formal representation. Consequently, in supposing that the items in the empirical relational system are beliefs and desires (from which he infers that their belief-like properties, such as having the contents that they do essentially, must be preserved by any system of formal representation) Crane (1995: p. 194) mischaracterizes a desideratum for a successful articulation of the measurement analogy.

7 Following the suggestion above concerning the empirical properties and relations to be preserved in representation by a formal relational system, a simplistic statement of such a theorem might take the following form. The empirical relational system (ERS) would have the form

<U, *tendstobringabout, tendstobebroughtaboutby*>

where U is a set comprising both internal states of the organism and states of affairs in the organism's environment, the latter conceived as the standing in a relation R of finitely many objects, $o_1, \ldots, o_n$. We shall refer to the state of affairs of relation R holding of $o_1, \ldots, o_n$ with the expression 'R$(o_1, \ldots, o_n)$'. 'Tendstobringabout' and 'tendstobebroughtaboutby' name relations that are borne between states both of the organism and the environment. (One is not the converse of the other: whereas the belief that P tends to be brought about by the state of affairs of P's obtaining, it is not the case that that state of affairs tends to be brought about by any belief.) We shall accordingly have such empirical relations to represent by means of the formal relational system as *tendstobringabout* (a, R$(o_1, \ldots, o_n)$) and *tendstobebroughtaboutby*(a, R$(o_1, \ldots, o_n)$), where the first relation is characteristic of desire and the latter of belief. The formal relational system (FRS) will have a domain comprising the set of real numbers R, as well as a set of Russellian Annotated Matrices, or RAMs, such that for every state of affairs R$(o_1, \ldots, o_n)$ in the ERS there is a RAM in the FRS having the n + 1-tuple $< o_1, \ldots, o_n, R >$ as its second element and as its first element a sentence S capable of expressing $< o_1, \ldots, o_n, R >$ in some context of utterance. The FRS will also contain the binary relations BEL and DES, defined on pairs whose first element is a real number and whose second element is a RAM. Among the relational facts that might hold in the FRS will thus be BEL(k, $<S, <o_1, \ldots, o_n, R>>$) and DES (k, $<S, <o_1, \ldots, o_n, R>>$), where k is a real number, S a sentence, and $<o_1, \ldots, o_n, R>$ a proposition that S expresses in some context of utterance.

With the aim of representing the ERS with the FRS, we begin with a mapping f that takes internal states into numbers, and states of affairs to RAMs. For the internal states, we use Axiom of Choice to ensure that the set of internal states has a well ordering, and we number all the states thus ordered starting with 1 and going upwards. To map propositions to RAMs, observe that there will in general be many RAMs having the same second element, and a function from a proposition in the ERS to the RAMs in the FRS must choose among these. Again, order all the RAMs having the same second element by means of Choice, and define f as follows:

> Where $R(o_1, \ldots, o_n)$ is a state of affairs in the domain of the ERS, define $f(R(o_1, \ldots, o_n,))$ as the first RAM according to the foregoing ordering whose second element is $<o_1, \ldots, o_n, R>$.

For a proof that FRS represents ERS according to our definition of f, it needs to be shown that as f has been defined.

> if *tendstobringabout* $(a, R(o_1, \ldots, o_n,))$, then $DES(f(a), f(R(o_1, \ldots, o_n,)))$ and

and

> if *tendstobebroughtaboutby* $(a, R(o_1, \ldots, o_n,))$, then $BEL(f(a), f(R(o_1, \ldots, o_n,)))$.

We need not attempt such a proof here.

8 Accordingly, it cannot be assumed without argument that different ways of representing mass are analogous to different systems of ascribing beliefs and desires. Unless it can be shown that beliefs and desires are the only internal states that can account for their possessors' optimizing within their environment, a less tendentious view is that different ways of representing mass are analogous to different ways of ascribing *attitudes*, which may or may not require to be described in terms of the concepts of belief and desire.

9 For a statement of this view see Foley, 1992, who in turn appeals to Locke. See also van Fraassen, 1989: p. 155.

10 More precisely, to each agent we ascribe a set X, which we shall call the *outcome space*, or set of all possible outcomes. The elements e_i of X are *atoms*. Set of elements of X are *events*. A *plausibility space for* δ is an ordered pair $(\Omega(X), P_\delta)$, where $\Omega(X)$ is an algebra on X and P_δ is a function from $\Omega(X)$ into the real numbers greater than or equal to zero. The union of all the atoms, which is also the top of the algebra, is the *certain event*, while the intersection of all the atoms, which is also the bottom of the algebra, is the *impossible event*. A *probability space for* δ is a plausibility space such that P_δ is a real valued function from $\Omega(X)$ to [0, 1] that satisfies the following axioms:

> P. $P_\delta(E) \geq 0$ for all $E \in \Omega(X)$.
> P. $P_\delta(X) = 1$, where X is the certain event.
> P. $P_\delta(E \vee F) = P(E) + P(F)$, where E and F are inconsistent elements of $\Omega(X)$.

11 The acts are assumed to be causally independent of the state of the world.
12 For more on this point see Harman, 1986: p. 22.
13 For further discussion of cases of this sort see van Fraassen, 1984 and Green and Hitchcock, 1994. Observe also that all propositions concerning past events may now have objective probabilities of either zero or 1. Even if we are all sure of this, we nevertheless – and without irrationality – have degrees of belief in many such propositions that are at neither extreme of the unit interval.
14 These points are discussed more fully in Stalnaker, 1984: pp. 90–1; Harman, 1986, Chapter 3; Maher, 1993, Chapter 6; and Kaplan, 1995. See also Maher, 1986, and Maher, 1993, Chapter 6, for an argument that degree of belief 1 in a proposition is neither necessary nor sufficient for belief in that proposition.
15 Kaplan, 1995, 1996, has argued that in some cases it can be rational to believe a proposition to which one assigns less than .5 degree of belief, but for the present discussion we need only maintain that it is possible to do so.
16 For further discussion of the relation of desire to degrees of desirability see the Introduction to Kyburg, 1984, and van Fraassen, 1989: p. 155. For Davidson's usage see, for instance, Davidson, 1963: p. 4, and 1978: p. 87. I speak only of Davidson's original account of pro-attitudes since in more recent work (for instance 1987: p. 41) he construes them as 'disposition[s] to act under specified conditions in specific ways'. Attaching a degree of desirability to a possible outcome of an available act seems to fit this characterization. Consequently, on this latter usage of 'pro-attitude', the decision theorist may accept that she imputes pro-attitudes to agents without being committed to the view that she imputes to them any of the posits of belief-desire psychology.
17 See, for instance, Nagel, 1970: p. 29, and Audi, 1973.
18 Schueler (1995: p. 34) develops this point.
19 For further discussion of the difference between belief and degree of belief in terms of the notion of aiming at truth, see Vickers, 1989. See also Collins, 1996 for an account of the connection between belief and being at risk of being mistaken.
20 We need not require certainty. Witness the 'trembling hand' hypothesis familiar to game theorists.
21 Her not taking P to be possible does not require her to believe P not to be possible. Also, her not taking P to be possible need not put her in violation of the decision-theoretic norm that one is to be certain of necessary truths, for what is necessary is relative to the agent's outcome space, and in the imagined situation the only necessary truth is that either the mathematical proposition or its negation holds.
22 See Tversky, 1975 for further discussion.
23 See Kahneman and Tversky, 1979, Tversky and Kahneman, 1992, and Wakker and Tversky, 1993 for formulations of prospect theory, which has been receiving attention in psychology, economics, political science, and linguistics. I shall discuss what has been called cumulative prospect theory.
24 In a recent formulation (Wakker and Tversky, 1993) the primitive concepts of the theory are these:

a finite set of S states of nature;
a finite set X of consequences, defined as containing 0 (the neutral outcome), and with all other elements of X construed as either gains or losses;
a function W from subsets of S into the reals satisfying $W(\{\}) = 0$; $W(S) = 1$; and $W(A) \geq W(B)$ whenever $A \supset B$.

These authors also formulate a decision rule different from that given in the text but too complex to admit of brief recapitulation here.

25 See, for instance, Slovic, Fischoff, and Lichtenstein, 1988 for evidence confirming prospect theory's prediction that outcomes that are merely probable are underweighted in comparison with outcomes that are obtained with certainty; and Tversky and Kahneman, 1992 for evidence confirming its predictions that agents are risk averse for gains and risk seeking for losses of high probability, and risk seeking for gains and risk averse for losses of low probability.

26 Although there is no need to argue here that prospect theory is superior to belief-desire psychology, we do well to observe that as a device for the explanation and prediction of action, belief-desire psychology is not unexceptionable. For instance, belief-desire systems of action explanation seem to have difficulty accounting for choices in which the agent appears not to be acting as if what she believes is true. For instance, a homeowner who believes that her house will not flood during the impending summer may nevertheless have good reason to purchase insurance against flood damage if she can do so at a reasonable price. She desires, let us suppose, to spend the quantity Q required to purchase such insurance on a vacation instead, and since she believes that her house will not flood, the belief-desire organon for understanding behaviour (which predicts that agents act as if what they believe is true) would seem to predict that the homeowner will spend Q on a vacation. Yet homeowners often, and with no lapse of rationality, choose insurance over a week at the beach. (For other difficulties with an account of the relation of belief, desire, and action according to which a person's beliefs are such that were they true her actions would satisfy her desires, see Vickers, 1976: Ch. 2.)

27 Maher (1993: pp. 5–8) argues that decision theory has the consequence that it is not always rational to calculate how to act, or what to think.

28 In spite of recognizing that decision theory does not always enjoin one to calculate how to frame a decision problem, Lance, 1995 asserts that Bayesian decision theory presupposes a concept of acceptance, and conjectures that other major forms of decision theory do as well. According to this author no matter how precisely the decision theorist specifies an agent's decision problem in the course of determining what it is rational for that agent to do, without presupposing a notion of acceptance it cannot be maintained that any agent in the same choice situation ought (prudentially) to do the same thing. As an example to motivate this thesis Lance considers a skater deciding whether to skate on a lake in light of the possibility of the ice being too thin. A decision matrix such as the following may seem appropriate:

	Thick ice	Thin ice
skate	+5	–20
don't skate	–1	–1

The trouble, according to Lance, is that this matrix fails to take into account the possibility of snipers in the woods determined to shoot anyone who skates without falling through the ice. In that case the outcome on the top left of the matrix would not be valued more highly than that on, say, the bottom right. One could, Lance goes on, attempt to take stock of the possibility of snipers by doubling the number of columns in the matrix, but this will just raise the issue of other factors the presence or absence of which is relevant to the assignment of values to outcomes. On pain of regress the theorist should instead take it that the agent just accepts that there are no snipers in the vicinity of the pond.

Introductory expositions of decision theory often remind the reader to ensure that the outcomes represented in any decision matrix be specified with sufficient precision for an appropriate assignment of values, and Lance has subtly failed to heed this admonition. The theorist's determination of the agent's degrees of belief, outcome space, and evaluations on elements of that outcome space is an empirical matter even when the theorist's concern is what the agent ought, prudentially, to do. If we are inclined to think that the evaluations in the above matrix are incorrect, that may indeed be due to the outcome space's being characterized too coarsely *by the decision theorist*. She may refine her characterization by, for instance, interrogating the agent, but once she has done this and developed a well-supported conception of the agent's outcome space and evaluations on elements thereof, there is simply no further work to do before calculating the rational action. We may continue to think that the agent should take into account further contingencies, such as the ice's turning to quicksand, but such a failing (if it is one) is beyond the decision theorist's purview in just the way that degrees of belief that satisfy the Kolmogorov axioms but that seem deviant in other respects are. It may be true that decision theory should ascribe acceptance states to agents (and at p. 164 Lance suggests that this is the upshot of his argument in spite of earlier (pp. 147–8) advertising his thesis as being that Bayesian decision theory presupposes a concept of acceptance), just as it may be that the theory should lay down stricter synchronic canons for degrees of belief than satisfaction of the Kolmogorov axioms, but our concern here is only what the theory does say.

29 In more recent work (1980b, 1985, 1989, 1990, and 1995) Davidson has pronounced a unification of radical interpretation and radical decision theory that he calls the Unified Theory of Meaning and Action. In formulating this approach he argues that subjective probabilities for various sentences (rather than the evidential base of 'holding true' various sentences) may be used to yield a theory of meaning for a speaker. These degrees of belief are to be elicited by the familiar technique of extracting cardinal utilities and subjective probabilities from preferences. (In Davidson's own case the preferences are between pairs of sentences.) The evidential base for the Unified Theory therefore need not depend upon attributions of belief, and so Davidson's more recent work supports my reply to his earlier argument.

30 For an account of assertion that attempts to elucidate the nature of the commitment that the assertor undertakes, see Brandom, 1983, 1994, Belnap and Green, 1993, and Green, 1995, 1999.

31 From the point of view of the decision theorist, assertion of 'I believe that P' is often just a way of asserting P, perhaps with an indication of tentativeness. On this account that assertion is often well glossed as 'P, I believe.' The verb phrase 'I believe' cannot, however, be semantically transparent in the way that 'It is true that' is, since sentences such as 'I believe that snow is white though snow is not white' are in spite of being pragmatically deviant not semantically inconsistent. The decision theorist may accommodate this datum by holding that when 'I believe' does make a semantic contribution to the sentence in which it occurs, the smallest indicative sentence of which it is a part is always a matter of fact (thought not in all possible words) false. (Green, 1997 discusses the parenthetical use of attitude verbs.)

32 Maher (1993: Ch. 7.1) has argued that Bayesian reconstructions of important episodes in the history of science inevitably smuggle in a concept of acceptance.

He discusses, among other studies, Dorling's (1974) reconstruction of Cavendish's reasoning concerning an experiment aimed at discerning the relation between the electrostatic force between charged particles and their distance. On Dorling's analysis the fact to be explained is that Cavendish's experiment and reasoning render it highly probable that the correct law is good approximation to the inverse-square law, but Maher objects that Dorling has explained an historical datum only if Cavendish did take it to be highly probable that the inverse-square law approximates the truth. The historical record only contains Cavendish's utterance, presumably with assertoric force, of:

> We may therefore conclude that the electric attraction and repulsion must be inversely as some power of the distance between that of the 2 + 1/50th and that of the 2–1/50th, and there is no reason to think that it differs at all from the inverse duplicate ratio.
>
> (Cavendish, 1879: pp. 111–12)

Maher contends that Dorling must assume that accepted propositions are given high probability, and shows this assumption to be false. Dorling, however, need only take it that the assertion of a proposition is good evidence that the speaker assigns it high probability. The speaker's making an unguarded assertion rather than explicitly registering her degree of belief may be due to any of a number of factors, not least of which would be the convenience of expression and the irrelevance (to the speaker's task) of registering that degree of belief. The evidence that an assertion provides may be defeated by, for instance, indications that the speaker refused certain bets on the proposition in question, but this fact does not vitiate an inference from Cavendish's speech act to his degree of belief. As a result Dorling's reconstruction of Cavendish's reasoning need not presuppose a notion of acceptance, much less an implausible thesis relating acceptance and personal probability.

33 The example is from Dennett (1991), who, while supporting the measurement analogy, construes both Brown and Jones as indispensably employing systems of belief and desire.

34 For such a rule within the belief-desire framework that applies to carefully circumscribed formal languages see Massey, 1992. Jackman, 1996 criticizes Wallace-style permutation rules.

35 This conclusion does not actually necessitate recourse to systems of the third sort. All that is needed is the premise that for every internal state represented as a belief (desire) in a system of the first or second kind, that representation has an image in some admissible transformation of the system of ascriptions of which it is a part in which the internal state is represented as a decision weight (degree of gain or loss). Consequently the wanted conclusion can be reached even if every empirically adequate system of propositional attitude ascription ascribes some beliefs or desires. Since systems of the third sort do seem empirically adequate, however, recourse to them has been made for the sake of expository simplicity.

References

Audi, R. (1973) 'The Concept of Wanting', *Philosophical Studies* 24: 1–21.

Belnap, N. and Green, M. (1994) 'Indeterminism and the Thin Red Line', in J. Tomberlin (ed.) *Philosophical Perspectives VII: Logic and the Philosophy of Language*, Atascadero, CA: Ridgeview Press, pp. 365–88.

Brandom, R. (1983) 'Asserting', *Noûs* 17: 637–50.
——(1994) *Making It Explicit*, Cambridge, MA: Harvard University Press.
Cavendish, H. (1879) *The Electrical Researches of the Honourable Henry Cavendish*, ed. J. Clerk Maxwell, Cambridge: Cambridge University Press.
Churchland, P. M. (1979) *Scientific Realism and the Plasticity of Mind*, Cambridge: Cambridge University Press.
——(1989) 'Folk Psychology and the Explanation of Human Behavior', in J. Tomberlin (ed.) *Philosophical Perspectives III: Philosophy of Mind and Action Theory*, Atascadero, CA: Ridgeview Press, pp. 225–42.
Collins, A. (1996) 'Moore's Paradox and Epistemic Risk', *The Philosophical Quarterly* 46: 308–19.
Crane, T. (1990) 'An Alleged Analogy between Numbers and Propositions', *Analysis* 50: 224–30.
——(1995) *The Mechanical Mind*, London: Penguin.
Davidson, D. (1963) 'Actions, Reasons, and Causes', repr. in Davidson, 1980a: pp. 3–21.
——(1974) 'Belief and the Basis of Meaning', repr. in Davidson, 1984: pp. 141–54.
——(1978) 'Intending', repr. in Davidson, 1980a: pp. 83–102.
——(1980a) *Essays on Actions and Events*, Oxford: Clarendon.
——(1980b) 'Towards a Unified Theory of Meaning and Action', *Grazer Philosophische Studien* 2: 1–12.
——(1984) *Inquiries into Truth and Interpretation*, Oxford: Clarendon.
——(1985) 'A New Basis for Decision Theory', *Theory and Decision* 18: 87–98.
——(1987) 'Problems in the Explanation of Action', in P. Pettit, et al. (eds) *Metaphysics and Morality: Essays in Honour of J. J. C. Smart*, Oxford: Blackwell, pp. 35–49.
——(1989) 'What is Present to the Mind?', in J. Brandl and W. Gombocz (eds) *The Mind of Donald Davidson*, Amsterdam: Rodopi, pp. 3–18.
——(1990) 'The Structure and Content of Truth', *Journal of Philosophy* 87: 279–328.
——(1995) 'Could There Be a Science of Rationality?', *International Journal of Philosophical Studies* 3: 1–16.
Dennett, D. (1987) *The Intentional Stance*, Cambridge, MA: MIT Press.
——(1991) 'Real Patterns', *Journal of Philosophy* 88: 27–51.
Dorling, J. (1974) Henry Cavendish's Deduction of the Electrostatic Inverse Square Law from the Result of a Single Experiment', *Studies in History and Philosophy of Science* 4: 327–48.
Dretske, F. (1988) *Explaining Behavior: Reasons in a World of Causes*, Cambridge, MA: MIT Press.
Field, H. (1981) Postscript to 'Mental Representation', in N. Block (ed.) *Readings in the Philosophy of Psychology*, Vol. II, Cambridge, MA: Harvard University Press, pp. 112–14.
Foley, R. (1990) 'Probabilism', in P. A. French et al. (eds) *Midwest Studies in Philosophy XV: The Philosophy of the Human Sciences*, Notre Dame, IN: University of Notre Dame Press, pp. 114–29.
——(1992) 'The Epistemology of Belief and the Epistemology of Degrees of Belief', *American Philosophical Quarterly* 29: 111–24.
——(1993) *Working Without a Net: A Study of Egocentric Epistemology*, Oxford: Oxford University Press.
Green, M. (1995) 'Quantity, Volubility, and Some Varieties of Discourse', *Linguistics and Philosophy* 18: 83–112.
——(1997) 'On the Autonomy of Linguistic Meaning', *Mind* 106: 217–43.

——(1999) 'Illocutions, Implicata, and What a Conversation Requires', *Pragmatics and Cognition* 7: 65–92.

Green, M. and Hitchcock, C. (1994) 'Reflections on Reflection: Van Fraassen on Belief', *Synthese* **98**: 297–324.

Harman, G. (1986) *Change in View*, Cambridge, MA: MIT Press.

Hunter, D. (1996) 'On the Relation between Categorical and Probabilistic Belief', *Noûs* 30: 75–98.

Jackman, H. (1996) 'Radical Interpretation and the Permutation Principle', *Erkenntnis* 44: 317–26.

Kahneman, D. and Tversky, A. (1979) 'Prospect Theory: An Analysis of Decisions Under Risk', *Econometrica* 47: 263–91.

Kaplan, M. (1995) 'Believing the Improbable', *Philosophical Studies* 77: 117–46.

——(1996) *Decision Theory as Philosophy*, Cambridge: Cambridge University Press.

Kyburg, H. (1984) *Theory and Measurement*, Cambridge: Cambridge University Press.

Lance, M. (1995) 'Subjective Probability and Acceptance', *Philosophical Studies* 77: 147–79.

Leeds, S. (1979) 'Church's Translation Argument', *Canadian Journal of Philosophy* 9: 43–51.

Maher, P. (1986) 'The Irrelevance of Belief to Rational Action', *Erkenntnis* 24: 363–84.

——(1993) *Betting on Theories*. Cambridge: Cambridge University Press.

Massey, G. (1992) 'The Indeterminacy of Translation: A Study in Philosophical Exegesis', *Philosophical Topics* 20: 317–45.

Matthews, R. (1994) 'The Measure of Mind', *Mind* 103: 131–46.

Nagel, T. (1970) *The Possibility of Altruism*, Oxford: Oxford University Press.

Richard, M. (1990) *Propositional Attitudes*, Cambridge: Cambridge University Press.

Schueler, F. (1995) *Desire*, Cambridge, MA: MIT. Press.

Slovic, P., Fischoff, B., and Lichtenstein, S. (1988) 'Response Mode, Framing, and Information-Processing Effects in Risk Assessment', in D. Bell, H. Raiffa, and A. Tversky (eds) *Decision Making: Descriptive, Normative, and Prescriptive Interactions*, Cambridge: Cambridge University Press, pp. 152–66.

Stalnaker, R. (1984) *Inquiry*, Cambridge, MA: MIT Press.

Suppes, P. and Zinnes, J. (1963) 'Basic Measurement Theory', in D. Luce et al. (eds) *Handbook of Mathematical Psychology*, New York: Wiley, pp. 1–76.

Swoyer, C. (1987) 'The Metaphysics of Measurement', in J. Forge (ed.) *Measurement, Realism, and Objectivity: Essays on Measurement in the Physical and Social Sciences*, Dordrecht: Kluwer, pp. 235–90.

Tversky, A. (1975) 'A Critique of Expected Utility Theory: Descriptive and Normative Considerations', *Erkenntnis* 9: 163–73.

Tversky, A. and Kahneman, D. (1992) 'Advances in Prospect Theory: Cumulative Representations of Uncertainty', *Journal of Risk and Uncertainty* 5: 297–323.

van Fraassen, B. (1984) 'Belief and the Will', *Journal of Philosophy* 81: 235–56.

——(1989) *Laws and Symmetry*, Oxford: Clarendon.

Vickers, J. (1976) *Belief and Probability*, Dordrecht: D. Reidel.

——(1989) 'An Agenda for Subjectivism', *Erkenntnis* 31: 397–416.

Wakker, P. and Tversky, A. (1993) 'An Axiomatization of Cumulative Prospect Theory', *Journal of Risk and Uncertainty* 7: 147–76.