PSYCHOLOGISM AND PSYCHOLOGY: THOMAS NAGEL'S THE LAST WORD*

It is frequently remarked that there was a common origin for the two philosophical traditions that we now call the analytical and the continental. This common origin lay in the rejection of the doctrine often called psychologism by, on the one hand, Gottlob Frege, the founder of analytical philosophy, and, on the other, Edmund Husserl, the founder of phenomenology (Dummett 1993). The historical myth among analytical philosophers is that Husserl was converted to his later anti-psychologism by Frege's rather brutal review of his 1891 book on the philosophy of arithmetic. But it is also often stressed that a general anti-psychologism was already in the air in the German-speaking world in the writings of Bolzano and others. Be that as it may, the fact is that the mature writings of Husserl share with Frege a belief that the proper study of thought should not be tainted by empirical psychology. The structure and inter-relation of thoughts cannot be understood by examining the particular transitions and operations that take place in individual minds. Both Husserl and Frege strove to understand the universal nature of thinking in a way that transcends the contingent peculiarities of individual thinkers or groups of thinkers.

Thomas Nagel's The Last Word is a contribution to this project of divorcing the stufy of thought from the study of thinking. Nagel's opponents are not those against whom Frege and Husserl battled. His target is the seemingly inexorable spread of subjectivism and relativism throughout the academy and indeed throughout philosophy itself. The Last Word offers a sturdy defence of the universality and objectivity of reason against these subjectivist and relativist opponents. Nagel tackles head on current forms of relativism in logic, language, science and ethics. His overall argument is that project of criticising reason is self-defeating. Every criticism of reason must be based on certain principles that cannot themselves be understood subjectively or relativistically. It is these principles that have the "last word" in any debate about the value and applicability of reason.

The Last Word can profitably be seen as an attack on a resurgence of psychologism. It is this aspect of the book that I will concentrate on in this critical notice. Nagel's book gives us an interesting perspective on an important concept, one that we need to make sense of if we are to understand not only the philosophical past but also the philosophical future. I shall distinguish several aspects of psychologism that Nagel tends to run together. The view I shall end up proposing is that the proper understanding of human reasoning depends on facts about human psychology in ways that are both more subtle than Nagel allows for and that invalidate some of his central claims and arguments. I shall discuss only the parts of the book that deal with reasoning and rationality.

Psychologism is a concept whose primary application is to discussions of rationality. The concept originated, of course, in the discussion of the most theoretical part of theoretical rationality – viz. deductive logic and the foundations of arithmetic. Here is a marvellous and justly famous passage from the Introduction to Frege's Basic Laws of Arithmetic. He is discussing the traps we can fall into by misunderstanding the concept of a logical law:

> The ambiguity of the word 'law' is fatal here. In one sense it states what is, in the other it prescribes what should be. Only in the latter sense can the laws of logic be called laws of thought, in laying down how one should think. . .But the expression 'law of thought' tempts us into viewing these laws as governing thinking in the same way as the laws of nature govern events in the external world. They can then be nothing other than psychological laws, since thinking is a mental process. And if logic were concerned with these psychological laws then it would be a part of psychology. And so it is in fact conceived. . . So in the end truth is reduced to the holding as true of individuals. In response I can only say: being true is quite different from being held as true, whether by one, or by many, or by all, and is in no way to be reduced to it. There is no contradiction in something being true which is held by everyone as false. I understand by logical laws not psychological laws of holding as true, but laws of being true. (Frege 1893 p.xv)

In this passage Frege draws a very sharp distinction between prescriptive principles of truth and descriptive laws of human reasoning. The prescriptive principles tell us how we ought to think, on the assumption that our thinking is guided by the norm of truth, while the descriptive laws merely inform us about how human beings do actually go about the business of reasoning. Any intrusion of psychological considerations into logic would, he thought, break down this divide between the prescriptive and the descriptive, compromising the normative force of the laws of logic. Hence the first of the three methodological principles that we find in

the Introduction to The Foundations of Arithmetic: "There must be a sharp separation of the psychological from the logical, of the subjective from the objective" (Frege 1884 p. x).

Nagel follows Frege in calling for just such a sharp separation of the psychological from the logical. In particular, he affirms the following three theses. The first thesis is effectively Frege's distinction between normative principles of reasoning and descriptive laws of thought:

> I. The fundamental principles of reasoning have normative force and make claims to universal validity

The second thesis is a mainstay of Nagel's attack on well-known attempts to undercut the validity of prescriptive principles by showing that they have their roots in purely contingent psychological facts, and hence cannot aspire to universal and unconditional validity.

> II. The fundamental principles of reasoning cannot be construed as the expression of contingent forms of life

There certainly seems, on the face of things, to be an important distinction to be made between two different claims that might be made in this area. It has certainly been argued that the apparent normative force of the fundamental principles of reasoning is merely a by-product of certain structures of oppression, or of the will to power, or of logocentrism, or whatever – and the conclusion drawn that the principles should be ignored and indeed resisted. But it has even more frequently been argued that certain aspects of reasoning will remain completely mysterious unless they are somehow grounded in contingent forms of life. Unlike the first, aggressive strategy, this is a defensive one. Its basic motivation, in essence, is the thought that there cannot be such a thing as Nagel takes reason to be. This thought is most easily motivated on metaphysical grounds (how can there be such things as universally valid principles with no naturalistic basis?), but it can also have an epistemological source (how could we have any insight into the nature of a universally grounded principle that is in no way grounded in our own nature or practice?). If we combine this thought, motivated in one or both of these two ways, with a desire to preserve some sort of prescriptive, normative framework for human reasoning then the defensive strategy seems an attractive option.

Nagel cannot accept this. As he puts it: "Being a realist about reason myself, I regard these reductive "rescues" as equivalent to skepticism; that is, they are forms of skepticism

about the reality of what I take reason to be" (p.9). That is to say, he effectively affirms that Thesis I entails Thesis II, and consequently that any denial of Thesis II is ipso facto a denial of Thesis I.

There is a third thesis to which Nagel subscribes that it is important to distinguish from the first two.

> III. The identification of fundamental principles of reasoning should be completely independent of psychological facts about how people reason.

Nagel allows (as indeed he must) that there may be first-order debate about whether some candidate or other really qualifies as a fundamental principle of reasoning. He doesn't give any examples, but quantum logic might be used to illustrate the point. Quantum logic denies a principle of inference that holds in classicical logic – the so-called distributive law that allows one to infer from "A & (B v C)" the disjunction "(A & B) v (A & C)". The motivation for this revision is the impossibility of simultaneously measuring the position and momentum of a particle. It may be true that a particle may has a given momentum and one or other of a range of possible positions at a given moment, but it cannot be true that a particle has one or other of a given set of position-momentum pairs at that moment. Nagel does not deny that such arguments can and should take place – and consequently he cannot deny that it is in some cases an open question what the fundamental principles of reasoning might be. But he certainly seems to be committed to denying that the resolution of such questions can appeal to psychological facts about human reasoning. In fact, it seems to follow from Thesis II. So, if Nagel is right that Thesis I entails Thesis II then Thesis III seems unavoidable.

It seems to me, however, that the inter-relations between these three theses are far more complex than Nagel is prepared to accept. Of course there are ways in which one might deny either or both Thesis II and Thesis III which will be incompatible with contintuing to affirm Thesis I. We have already seen how Thesis II can be denied as part of a radical strategy explicitly aimed at underminding Thesis I. It should be clear also that Thesis III can be incompatible with Thesis I. One might deny Thesis III by maintaining, as Herbert Simon and others have done (Simon 1982), that no normative theory of rationality can permit too great a discrepancy between the normative and the descriptive, and consequently call for the

construction of a theory of bounded rationality reflecting, not just widespread logical flabbiness, but also the short-cuts, heuristics and "quick and dirty" strategies that people tend to use in decision-making. According to Simon and others, the principle of "ought" implies "can" requires us to effect a radical down-shift in the demands that we place on rational thinking and rational decision-making. Our normative standards must be tailored to the actual reasoning habits of normal human subjects. Now, whatever the merits of Simon's proposal, it seems clear that it threatens Thesis I. The prescriptive norms look in danger of becoming conditional norms – norms that are conditional on people continuing to reason in the way that they actually do. Any significant change in reasoning practices seems to require a change in norms. And this surely cannot be acceptable to a believer in the universal validity of prescriptive norms of rationality.

But what I want to suggest is that, although denials of Theses II and III may involve or entail repudiations of Thesis I, this is not necessarily the case. There are ways of denying the second and third theses that are compatible with an uncompromising endorsement of the first. I will start with Thesis II. Here it is again:

> II. The fundamental principles of reasoning cannot be construed as the expression of contingent forms of life

I borrow the phrase "expression of contingent forms of life" from Nagel himself. The passage in which it appears is revealing and worth quoting:

> However much one may try to construe one's concepts and thoughts naturalistically, as the expression of contingent forms of life, the logic of the effort will always generate thoughts too fundamental for this, thoughts which one cannot get outside of in this way and which one must employ in trying to view from the outside anything else that one does. (pp.32-33)

These thoughts, rather than any facts about where they came from or who thinks them, have what Nagel calls the last word. They are the ultimate court of appeal.

It is important in understanding Nagel's project to understand what he means by saying that a given principle of reasoning is the last word. We find a clue in the following passage, where he sketches out a line of argument that makes many appearances in the book:

> It is not just that in criticizing each part of our system of beliefs we must rely on the rest, whatever it may be. The thoughts that enter into such criticism must aspire to a universality that is lacking in the thoughts criticized. Some of these

> standards have to be discovered; others are simply those basic and inescapable forms of reasoning that enter into all possible criticism – even when some examples of them are among the objects of criticism. The serious attempt to identify what is subjective and particular, or relative and communal, in one's outook leads inevitably to the objective and universal. (p. 16)

Since this line of argument effectively maintains the unavoidability of depending upon certain objective principles of reason when one seeks to justify one's practice, let us call it the argument from unavoidability

It seems to me that there are two possible lines of response to this argument, a weak response and a strong response. The weak response drives a wedge between two concepts that Nagel runs together – the concepts of understanding and justification. There is no need to follow Nagel in thinking that understanding and justification will run out at the same time. The weak response to Nagel accepts the conclusion of the argument from unavoidability. That is, it accepts that the process of justification will inevitably come to a halt with principles whose validity is independent of our point of view. But, nonetheless, it still seeks to deny Thesis II, on the grounds that, even though the process of justification must inevitably come to an end with certain ground-level principles (the principles that are, in Nagel's phrase, the last word), the process of understanding cannot stop there. Those principles themselves need to be explained. They cannot be self-standing, even though there is a sense in which they must be self-justifying.

It is certainly a shortcoming of Nagel's discussion that he neglects the distinction between justification and explanation. The distinction can easily be seen from discussions of perceptual knowledge. It seems clear that the justification of perceptual beliefs about what is the case lies in basic perceptual states – in how the world appears to the senses. Such states are, normally (that is to say, when one has no reason to think that all is not as it should be), the last word. They are where justification stops. But the process of understanding does not stop at the same point. It is perfectly legitimate to ask why the world should appear to the senses the way it does. The process of understanding will come to an end with contingent psychological facts about the specific constitution of our sensory apparatus – as opposed, for example, to the sensory apparatus of a bat or a rat. But to say that our perceptual states are grounded in such

contingent psychological facts is not to say that these contingent psychological facts are the ultimate court of appeal, the final justifiers. That would require a further step – and it is certainly optional.

It seems, then, that Nagel's unavoidability argument does not entail Thesis II. The conclusion of the unavoidability argument can be reconciled with the denial of Thesis II by driving a wedge between the concepts of justification and understanding. We might see our reasoning practices as the expression of contingent forms of life in the same way as we see our perceptual systems as the expression of contingent forms of life. Even though they are the last word in matters of justification, they can and should be understood in terms of contingent facts about our forms of life.

On reflection, however, applying the distinction between understanding and justification does not have the same immediate justification when applied to normative principles of reasoning that it does when applied to the relation between perceptual experiences and perceptual judgments. One wonders what it might even mean to say that our normative principles of reasoning should be understood as the expression of a contingent form of life. In the case of perceptual systems the matter seems clear enough. Our perceptual systems are the expression of a contingent form of life because there is a quite clearly defined sense in which they could have been otherwise. Quite apart from the distant possibility that we could now be perceiving the world via echolocation, it is clear that we might have had a visual system sensitive to a different range of properties or a far more acute olfactory system. A "contingent form of life" is one that might have been otherwise. But it is far from clear how our normative principles of reasoning could be understood in a way that brings out a way in which they could have been otherwise.

Of course, it is possible that we could have found different inferential principles self-evidently valid – just as it is presumably the case that we could have found the basic principles of some non-Euclidean geometry more intuitively obvious than Euclid's parallel postulate. But these would just be descriptive facts about how we reason – not evidence that the normative principles governing our reason had changed. This will not tell us what it would be like for a

different set of normative principles to govern our thinking. One might think that the normative force of the principles governing deductive reasoning derives from the imperative to aim for truth in our beliefs, and consequently for the preservation of truth in our inferences. The requirements of deductive rationality are hypothetical imperatives. If the relevant categorical imperative were different (if, for example, the imperative to aim for truth were replaced by the imperative to aim for beliefs that promote inclusive fitness – or, for that matter, the will-to-power), then we might see a different set of normative principles. That is to say, there might be circumstances in which fitness was promoted more by committing the fallacy of denying the antecedent than by the valid principle of denying the consequent (Nagel's favourite example of an inescapable principle of reasoning). Would this not be a case in which different principles of inference had normative force? One hesitates to say that it would be, simply because it seems impossible to explain the notion of deductive inference without appeal to the preservation of truth. In any case, Nagel is likely to say that there is a deep incoherence in any such proposal. Either the proposal rests upon the true claim that regular employment of the fallacy of denying the antecedent will maximize inclusive fitness or it does not. If it does not then it is worthless. But if it does then it is hard to see how inclusive fitness could be promoted by not restricting oneself to drawing from it only further true principles. This would be a higher-order version of the unavoidability argument.

Of course, the charge of incoherence is much harder to make stick than Nagel appears to recognize. In this case it clearly assumes that the revisionary reasoner is enaging in second-order reflection on what justifies his first-order principles. And there are ways of describing the hypothetical situation that do not allow this assumption to get a grip. One might imagine a community of purely first-order reasoners who merely make all and only those inferential transitions which they find self-evident. They neither reflect, nor are capable of reflecting, on how those principles fit together or why they should find them self-evident. It is hard to see how the unavoidability argument could get a grip. The unavoidability argument falls into the general category of anti-skeptical arguments that set out to show that, as soon as the skeptic sets out to defend his position, he will inevitably adopt precisely the views and principles he is criticising. But such anti-skeptical arguments only have force against skeptics who try to

defend their choice of deviant reasoning practices – and there seems no reason why we should only consider such skeptics when considering the possibility of different norms of reasoning.

A defender of Nagel will at this point no doubt say that the contrived nature of the example effectively proves his point. The imperative to have true beliefs and consequently to preserve truth in inferences cannot be merely a contingent part of our form of life. The norm of truth imposes itself as soon as we start to reflect upon our reasoning practices, and such reflection is inevitable in anything that is to count as a cognitive community.

The debate seems deadlocked. Let us move on then to the strong response to the unavoidability argument. Recall that the weak response is essentially a two-pronged strategy. It distinguishes the process of understanding from the process of justification, maintaining that the process of understanding can continue even when the process of justification has come to an end. It is not in justifying but in explaining our fundamental practices of reasoning that we fix their origins in our contingent form of life. In contrast, the strong response does not draw a comparable distinction between explanation and justification – or at least it does not for the prescriptive principles of reasoning. The strong response quite simply denies the claim that runs through Nagel's book to the effect that there can be no justification of the fundamental principles of deductive reasoning – the simple laws of logic are the last word.

The question of whether the basic principles of deductive reasoning can be justified is, of course, the inverse of the question of whether they can effectively be criticized. The process of justification is intended either as a response to criticism or to forestall potential criticism. Nagel's confidence that the basic principles of deductive reasoning cannot be justified is founded in his conviction that they cannot effectively be criticized. This in turn rests upon the unavoidability argument in various of its manifestations – effectively, we cannot criticize the laws of logic without self-defeatingly employing precisely the laws we are trying to criticize. Nagel does not take into account, however, that the laws of logic can be criticized in different ways. He restricts himself to rebutting a rather implausible imaginary figure, namely, a complete skeptic about the laws of logic who is nonetheless prepared to argue in defense of his views. It is clear that the sort of justification that would satisfy such an opponent cannot be

given – for more or less the same reasons that the opponent's position cannot be coherently formulated. But other, weaker types of justification may nonetheless both be possible and required in the face of different sorts of criticism.

The really interesting issues about the justification of logical laws arise not when one asks whether the practice of deductive reasoning as a whole can be justified, but rather when one asks which logical system correctly represents the principles of deductive reasoning. The project of finding a "correct" logical system is partly the project of making explicit our actual inferential practices. This, in turn, can be viewed as a matter of striking a balance between intuitive assessments of the validity or invalidity of ordinary language arguments and the assessments of validity that emerge when those arguments are represented in formal calculi in various ways. Given a proposed logical system one might criticize it either for counting as valid inferences which intuitively seem invalid, or for failing to accommodate inferences that seem intuitively valid – or alternatively one might criticize it for failing to accommodate entire patterns of reasoning upon which we pre-theoretically rely. The debates provoked by any of these three lines of criticism will have no bearing on Thesis II. To say that judgments about the correctness of a logical system are answerable to intuitive assessments of the validity of arguments is not to ground the principles of reasoning in the contingencies of our form of life. Nagel makes this point very clearly. To judge that an argument is valid is to judge it impossible that the premises of that argument be true and the conclusion false, and

> The judgment that it is impossible or inconceivable that the premises of a proof be true and the conclusion false relies on our capacities and incapacities to conceive of different possibilities, but it is not a judgment about those capacities, and its object is not something that depends on them. (p.57)

This is quite right, but the process of finding a correct logical system is not simply a matter of matching up formally valid inferences with intuitively valid inferences. It can involve revising our conception both of which inferences are formally valid and of which inferences are intuitively valid. We have already seen one example of this in the proposed revision of the distributive law in quantum logic. For present purposes a more relevant example is Michael Dummett's version of the intuitionistic challenge to classical logic and the law of excluded middle – that is to say, the status of 'p v ~p' as a logical truth.

As is well-known, Dummett's central claim is that the debate between classical logic and intuitionistic logic should be reconstrued as a debate about meaning and the correlative notion of understanding (Dummett 1991). That is to say, the decision as to which logic is the correct one is based on arguments concerning the appropriate meanings for the logical constants. The intuitionistic argument is that classical logic accords certain logical constants a meaning that cannot be understood – and hence endorses argument forms which it should not (even though they might intuitively appear to be valid, indeed paradigms of validity). The logical rules that are invalidated according to intuitionistic propositional logic include the Law of Excluded Middle ('p v ~p'); the law of Double Negation (that from '~~p' one can conclude 'p'); Classical Reductio (that if 'p' leads to a contradiction one can conclude '~p'; and what is sometimes called the rule of Nonconstructive Dilemma (that if both 'p' and '~p' entail 'q' one is entitled to conclude the truth of 'q').

The general strictures concerning meaning that govern Dummett's theory are derived from an attack on the sort of crude psychologism that Frege. The meaning of any linguistic expression is what is communicated between individuals. It is what both speaker and hearer understand in instances of successful communication. If the meaning of a linguistic expression is exhausted, as Dummett thinks it is, by its role in communication then it follows that there can be nothing private about meaning: "An individual cannot communicate what he cannot be observed to communicate" (Dummett 1978 p.216). So, the knowledge that we have of a linguistic expression is knowledge that must be capable of being manifested in use. It would be natural to hold that, when the linguistic expression is a sentence, the appropriate manifestation would consist in the ability to recognize it as obtaining when it does – or, more realistically, the understanding of some sort of procedure whereby one can put oneself in a position to determine whether or not it does obtain.

Dummett's basic argument is that there are many types of sentence for which this type of manifestation of understanding is in principle unavailable. Undecidable sentences in mathematics provide an obvious example – Goldbach's conjecture (that every even number greater than 2 is the sum of two primes), for example, or the Prime Pairs conjecture (that there

are infinitely many pairs of prime numbers (p, q) such that q = p + 2. But so too do a range of non-mathematical sentences: sentences about inaccessible regions of space-time; certain unverifiable subjunctive conditionals (e.g. the conditional that some long-dead person who was never in a situation of danger would have acted cowardly had they found themselves in such a situation); or sentences about another person's sensations. In none of these cases (according to Dummett) can our grasp of the meaning of a sentence be manifested in the ability to recognize its truth. All that can be manifested is our ability to recognize evidence that licenses us, to some degree or other, in asserting the sentence in question.

The conclusion Dummett draws is that the notion of meaning needs to be divorced from the notion of truth. We should replace the notion of truth with the weaker notion of warranted assertibility (although the two notions will coincide for those sentences that are decidable). This has obvious and immediate consequences for the basic principles of deductive reasoning. If the meaning of a sentence is to be given in terms of its assertibility conditions (rather than its truth conditions) then the meaning of the Law of Excluded Middle is not 'Either p is true or p is not true' but 'Either p is warrantedly assertible or its denial is warrantedly assertible'. And once the Law of Excluded Middle is understood in these terms then it is clear that it fails to hold for many well-formed and grammatical sentences. But without the Law of Excluded Middle then the argument forms identified earlier cease to be universally valid. It is clear that the LEM is presupposed by Double Negation, Classical Reduction and Nonconstructive Dilemma. So they have no place in the "correct" logic – even though they appear to have a central place in informal reasoning.

Now, I don't want to take a stand on this line of argument, although it strikes me as plausible in the extreme. The point I want to extract is a methodological one. There is a sense in which it is a form of psychologism, because the range of what is assertible is fixed by certain psychological contingencies. These are contingencies about what we are and are not in a position to assert. It is true that, in the case of anti-realism about mathematics, assertibility is tied to a (more or less) formally determinate notion, viz. effective decidability. But the problematic domain is not of course confined to mathematics.

It might seem to be stretching a point to call this psychologism – particularly in view of Dummett's own relentless and Frege-inspired opposition to psychologism (as manifested, of course, in the manifestation requirement). But the important point is that it shows the importance of separating out two things that Nagel does not keep firmly distinguished and reveals a position in logical space that he does not take into account. It is possible that the content of logical laws might be fixed in a way that depends upon certain features of our contingent form of life and yet that does not compromise the normative force of those laws. This is particularly significant when one bears in mind how Nagel describes the basic principles with which justification stops:

> The last word, with respect to such beliefs, belongs to the content of the thought itself rather than to anything that can be said about it (p.64)

This may indeed be so - although one may feel with respect to the laws of logic that Nagel is underestimating the potential significance of proofs of soundness and completeness. But Nagel's comments about the content of the belief having the last word are not incompatible with perhaps the most plausible way in which logic might be construed as the expression of a contingent form of life. The content of a logical belief might have the last word even though we need to understand the form of life within which that belief is embedded if we are properly to understand what that content is. Alternatively, psychologism at the level of content need not compromise normativity.

Let us go back, then, to our three principles

> I. The fundamental principles of reasoning have normative force and make claims to universal validity.
>
> II. The fundamental principles of reasoning cannot be construed as the expression of contingent forms of life.
>
> III. The identification of fundamental principles of reasoning should be completely independent of psychological facts about how people reason.

The conclusion to draw from the discussion of deductive reasoning is that, pace Nagel, Thesis I does not entail Thesis II. It is possible to deny Thesis II without compromising the normativity and universal validity of the fundamental principles of deductive reasoning. Let us turn now to Thesis III. Despite a superficial similarity, Thesis III is very different from Thesis II. The line of reasoning we have discussed as a possible objection to Thesis II cannot serve as

an objection to Thesis III. The psychological facts we have been discussing are psychological facts about understanding that constrain the interpretation of logical constants. They are not psychological facts about how people reason. It follows, therefore, that one can deny Thesis II without ipso facto denying Thesis III. What are there independent reasons for rejecting Thesis III?

Nagel considers and rejects one possible set of grounds for rejecting Thesis III. He points to the well-known experimental evidence suggesting that even sophisticated reasoners have inordinate difficulty with applying modus tollendo tollens. This, despite what Nagel repeatedly says, is not the same as the argument form of contraposition. Contraposition is an inference that takes one from one material conditional 'P Q' to another '~Q ~P', whereas modus tollendo tollens is the inference from a material conditional 'P Q' and the denial of its consequent '~Q' to the denial of its antecedent '~P'. As it happens, people cannot apply either of them very well in ordinary everyday reasoning. But, as Nagel quite rightly points out, this has no bearing on their prescriptive force. Prescriptive principles of deductive reasoning cannot be accountable to psychology in this crude sense. The normativity of a logical law is not parasitic on its widespread acceptance.

In part this is because the laws of deductive logic have, as we have already discussed, an over-arching justification in terms of truth-preservation. Since it is impossible that applying modus tollendo tollens should ever take us from truth to falsity, its place as a normative principle of reason seems secure. The same holds true of contraposition – as long, that is, as we take it to be a principle governing the behaviour of the material conditional. As it happens there are ordinary language instances of invalid instances of what appears to be contraposition – but the conclusion to draw from these, of course, is that the "If. . . then. . ." of ordinary language is not the ' ' of the propositional calculus. There are interesting and important issues about the formal analogues of the various conditionals of ordinary language, and consequently about the appropriate formal theory to apply in assessing ordinary language argumentation, but I mention them merely to pass swiftly by.

It would be too quick to conclude that the role played in deductive reasoning by the requirement of truth-preservation completely rules out the possibility that the normative principles of deductive logic might be somehow responsive to psychological facts about how people reason. After all, it might be the case that psychological facts about how people reason lead us to revise the requirement of truth-preservation. It is true that no psychological fact is ever going to lead us to accept a normative principle that fails to be truth-preserving. But psychological facts about how people reason might lead us to a stricter conception of validity, for which simple truth-preservation remains a necessary but is no longer a sufficient condition. This seems, in fact, to be one of the motivations behind relevance logics (Read 1988). The characteristic feature of relevant logics, in brief, is the demand that a valid argument from premises $A_1 . . . A_n$ to conclusion B be not simply truth-preserving, but also involve deducing B from $A_1 . . . A_n$ where genuine deducibility involves the actual employment of the premises in the derivation of the conclusion (and hence the relevance of the premises to the conclusion). Relevance logics have the characteristic of denying the validity of the argument form of disjunctive syllogism (modus tollendo ponens). It is no longer possible to infer 'Q' from 'P v Q' and '~P'. In particular disjunctive syllogism fails when the original disjunction is derived from the principle sometimes known as addition – namely, that when one has established 'P' one can infer 'P v Q' for any arbitrary 'Q'. One interesting corollary is the consequent invalidity of the argument form Ex Falso Quodlibet which allows one to derive any conclusion from a contradiction. Be that as it may, the point for present purposes is that the relevance logician's critique of classical logic is at least partly motivated by a desire to be faithful to the way in which we actually reason.

Whether or not one accepts this account of the motivation for relevance logic, once we move outside the realm of deductive logic no such over-arching constraint applies. There are no benchmarks for non-deductive reasoning that play a role comparable to that of the requirement of truth-preservation. This allows psychological facts about reasoning practices to get a grip in a way that they cannot in deductive reasoning. Or rather, the appeal to psychological facts about reasoning practices is not immediately blocked in the way it is in deductive reasoning. But how might such an appeal be brought to bear? Some well-known

criticisms of classical decision theory will illustrate. The normative maxim of classical decision theory is the injunction to adopt the course of action that maximizes expected utility, where the expected utility of a course of action is given by the sum of the products of multiplying the probability of each of its possible outcomes by that outcome's desirability. Every decision situation involves a choice between a range of feasible courses of action. Each of those feasible courses of action will have a range of possible outcomes, which can be assigned numerical probabilities and desirabilities. The expected utility of each possible outcome is calculated by multiplying its probability by its desirability. The expected utility of a course of action is obtained by summing the expected utilities of its possible outcomes. The injunction to maximize expected utility is obeyed by choosing the course of action with the highest expected utility.

The normative principle that one should maximize expected utility has great elegance and simplicity. But it has also been frequently criticized. One well-known criticism was raised by Maurice Allais (1953). His criticism in essence is that the principle of maximising expected utility is too narrow. Utility is not the only sort of value. What is interesting about it is that he seems to be arguing from descriptive premises to a critique of a normative principle. He considered two decision situations. In each decision situation one is offered a choice between two gambles, each with three possible outcomes. The possible outcomes for these gambles have the same probabilities in each lottery and in both decision situations, but the pay-offs vary.[1]

In the first decision situation the choice is between a lottery (Gamble A) in which every ticket will win a particular sum, say 500,000 Euros, and another lottery (Gamble B) on which one has a 10% chance of winning a subtantially greater sum, say 2,500,000 Euros, an 89% chance of winning the 500,000 Euros, but a 1% chance of not getting anything at all. Confronted with this type of decision situation people overwhelmingly choose the first option (Gamble A) - and one can see their point, given that gamble A offers 500,000 Euros as a dead certainty.

In the second decision situation the choice is between a lottery (Gamble C) in which there is an 11% chance of winning 500,000 Euros, but an 89% chance of getting nothing, and a second lottery in which the chance of getting nothing is increased slightly to 90%, but with the remaining 10% of tickets producing a return of 2,500,000 Euros. Allais noted that the same people who chose gamble A over gamble B will also tend to choose gamble D over gamble C. Here too one can see why. A 1/10 crack at 2,500,000 Euros seems better than an 11/100 chance of 500,000. There's a marginally greater chance of winning with gamble C, but the pay-off is five times greater. Surely it's better to take the risk.

The problem is that that this pattern of choices runs directly counter to the prescriptions of the expected utility principle.[2] So, whatever people are doing when they chose A over B and D over C it doesn't look as if they can be maximising expected utility. This poses a challenge to the prescriptive nature of the principle of maximising expected utility in a way that the regrettable tendency of otherwise intelligent people to fail to apply modus tollens does not pose a challenge to the status of modus tollens as a prescriptive principle. Why not? The simple answer is that the normative status of expected utility theory cannot be dissociated from its predictive power in the way that the normative status of deductive logic can. The reason for this emerges when one considers the most obvious objection that can be raised to the principle of maximising expected utility, namely, that it is empty because people just do not and cannot assign determinate and numerically precise probability and utility values to the different anticipated outcomes of given courses of action. There is no point having a normative principle governing how one should comport oneself relative to assignments that one does not and cannot have. The obvious objection has an equally obvious answer, namely, that expected utility theory is an idealization with a solid basis. The solid basis derives from the existence of formal techniques that allow one to work out what an individual's probability and utility assignments are. These are the so-called representation theorems which show how to map, for example, an agent's choices between different gambles onto two functions, one which can plausibly be taken to represent degrees of belief (assignments of subjective probability) and the other to represent desirabilities or assignments of utility (Ramsey 1931, Savage 1954). Of course, not just any set of choices can be mapped back into a probability function and a utility

function. The choices have to respect some basic constraints. They have to be transitive, for example. If you choose a over b and b over c then you'd better not prefer c over a. They also have to be connected – given any two options one must either choose one over the other or be indifferent between them. Most significantly, for present purposes, they have to respect what is called the sure-thing principle. This is the principle that one's choice between two different gambles, or two different courses of action, should not not depend upon outcomes that are constant across both gambles or courses of action. So, if you have a choice between two routes you can take to walk to the station to catch a train and you know that each route involves a 1% chance of slipping on a banana skin and spraining your ankle, then that shouldn't be a factor in your decision (assuming, of course, that staying at home isn't an option).

The Allais paradox highlights a weakness in the sure-thing principle. What grounds could one have for preferring Gamble A over Gamble B in the first decision situation? Obviously, it's got nothing to do with the outcome that is constant in the two cases, namely the 89% chance that one will win 500,000 tickets. So it must that one attaches greater expected utility to the fact that in gamble A the remaining 11% of the tickets will also produce a return of 500,000 while in Gamble B 10% of the tickets will produce a return of 2,500,000 while 1% will leave the gambler empty-handed. What about the second decision situation? Well, one's choice here can't be determined by the 89% chance of getting nothing, if the sure-thing principle is to be believed. So it must be determined by the rewards allocated to the other 11% of tickets. But this can't be right, because the pay-offs for the remaining tickets are exactly the same in the two decision situations. So, if one chooses Gamble A in Decision I then one is committed to choosing Gamble C in Decision II. But this isn't what happens.

The implications should be clear. If the patterns of choice displayed in the Allais paradox are indeed diagnostic of how people reason then the normative status of expected utility theory, at least in its classical form, is in trouble. Numerically determinate assignments of probability and utility are theoretical entities that explain observable choices and preferences. There is no content to a normative principle of practical rationality defined over theoretical

entities that do not appear capable of doing the explanatory work for which they were introduced. But the Allais paradox shows us that our only route to the theoretical entities concerned involves assumptions about choice and preference that are both empirical and empirically unacceptable.

Of course, I am not offering this as a knock-down argument against the normative principle of maximizing expected utility. On the contrary, I am convinced that some version of the principle must be normatively valid – just not the one that we find embedded in classical decision theory. The point I want to make is a methodological one. The process of finding a correct version of the expected utility principle will have to be sensitive to the psychological facts about reasoning that emerge from the Allais and other related empirical studies (Ellsberg 1961, Kahnemann and Tversky 1979). If this is right then what I earlier identified as Thesis III is false. The identification of fundamental principles of reasoning cannot be completely independent of psychological facts about how people reason.

* * *

Let us look again at the three theses that we have been discussing.

> I. The fundamental principles of reasoning have normative force and make claims to universal validity.
>
> II. The fundamental principles of reasoning cannot be construed as the expression of contingent forms of life.
>
> III. The identification of fundamental principles of reasoning should be completely independent of psychological facts about how people reason.

Nagel, like Frege, not only endorses all three but thinks that important relations of entailment hold between them. He maintains that one cannot retain Thesis I if one does not defend Theses II and III. That is to say, the psychology of reasoning cannot have any bearing on how we understand the fundamental normative principles of reasoning – on pain of opening the door to the forces of darkness. I have tried to suggest, however, that this is overly pessimistic. There are ways of denying both Thesis II and Thesis III that leave Thesis I untouched and keep the forces of darkness at bay. Or, in other words, attention to the psychology of reasoning does not lead inevitably to the excesses of psychologism.[3]

Bibliography

Allais, M. 1953/1979. Criticism of the postulates and axioms of the American school. Reprinted in Moser (ed.) 1990.

Dummett, M. 1991. The Logical Basis of Metaphysics. London. Duckworth.

Dummett, M. 1993. The Origins of Analytical Philosophy. London. Duckworth.

Ellsberg, D. 1961. Risk, ambiguity and the savage axioms. Reprinted in Gärdenfors and Sahlin 1988.

Frege, G. 1884. Die Grundlagen der Arithmetik. Translated by J. L. Austin, The Foundations of Arithmetic. Oxford. Basil Blackwell. 1950.

Frege, G. 1893. Grundgesetze der Arithmetik. Partially translated by M. Furth, The Basic Laws of Arithmetic. Berkeley. California University Press. 1964.

Gärdenfors, P. and Sahlin, N.-E. 1988. Decision, Probability and Utility. Cambridge. Cambridge University Press.

Moser, P. K. 1990. Rationality in Action: Contemporary Approaches. Cambridge. Cambridge University Press.

Nagel, T. 1997. The Last Word. Oxford. Oxford University Press.

Ramsey, F. 1931. Truth and Probability. Reprinted in Gärdenfors and Sahlin 1988.

Read, S. 1988. Relevant Logic. Oxford. Basil Blackwell.

Savage, L. 1954. The Foundations of Statistics. New York. John Wiley.

Simon, H. 1982. Models of Bounded Rationality. Cambridge MA. MIT Press.

---

[*] Thomas Nagel, The Last Word, New York/Oxford: Oxford University Press, 1997, ix + 147 pp., ISBN 0-19-5108345, £16.99).

[1] The pay-off table is:

	Outcome 1 p (1)= 0.01	Outcome 2 p (2)= 0.10	Outcome 3 p (3)= 0.89
DECISION I			

Gamble A	500,000 Euros	500,000 Euros	500,000 Euros
Gamble B	0 Euros	2,500,000 Euros	500,000 Euros
DECISION II			
Gamble C	500,000 Euros	500,000 Euros	0 Euros
Gamble D	0 Euros	2,500,000 Euros	0 Euros

[2] This can be easily appreciated. The expected utility calculation in Decision I can be represented as follows (leaving out the outcomes with zero pay-offs/probability):

EU (DECISION I) 1 U(500,000 Euros) > 0.1 U(2,500,000 Euros) + 0.89 U(500,000 Euros)

Simplifying by subtracting the final figure from both sides we get

EU (DECISION I) 0.11 U(500,000 Euros) > 0.1 U (2,500,000)

That is to say, the expected utility of 500,000 Euros with probability 1 is valued more highly than the sum of 2,500,000 Euros with probability 0.01 and 500,000 Euros with probability 0.89. But the expected utility calculation in Decision II can be represented as follows

EU (DECISION II) 0.1 U(2,500,000 Euros) > 0.10 U(500,000 Euros) + 0.1 (500,000 Euros)

or, more simply,

EU (DECISION II) 0.1 U(2,500,000 Euros) > 0.11 U(500,000 Euros)

This straightforwardly contradicts the putative expected utility calculation for the first decision situation.

[3] This paper was written for a conference on The Last Word held on May 15 1999 at the Centre for Hermeneutics and Analytic Philosophy at Anglia Polytechnic University, UK. Thanks to Neil Gascoigne and Jonathan Derbyshire for organizing an enjoyable occasion. Thanks also to my colleagues at Stirling for allowing me to take a year's leave to hold a visiting appointment at CREA, Ecole Polytechnique, Paris. This paper was written during my tenure of that appointment.