

Logic, Reasoning and the Logical Constants

PASCAL ENGEL
*Université Paris Sorbonne
and Institut Jean Nicod*

Introduction

Consider the three questions:

- (a) how do we reason?
- (b) how ought we to reason?
- (c) what justifies the way we ought to reason?

Question (a) is a descriptive one: it calls for an account, presumably a psychological one, of how humans reason, a theory of their reasoning abilities and of their reasoning performances. Question (b) is normative: it calls for a theory of how we evaluate whether a reasoning is good or bad, correct or incorrect, valid or invalid. Usually this role is devoted to logic. Question (c) is, so to say, meta-normative; it calls for the justification of our norms and standards of reasoning, possibly against other norms.

On many accounts, the answers to these three questions differ widely, and on many accounts they must do so. On the classical view of logic as a normative discipline, there is a strong distinction between the way we reason actually and the way we should reason, and the attempt to derive the latter from the former, is considered as the most obnoxious fallacy about the nature of logic: psychologism. Today's psychologists of reasoning, however, guard themselves against this fallacy. They are all the more suspicious of it that they find that there is a large gulf between our actual reasoning practices and logic as the theory of valid reason-

ing. So both on the classical philosophical view and on the contemporary view in psychology, (a) and (b) are independent. It is not clear either than question (c) is related to the others. What justifies our logical norms is not related to the way we actually reason, for it would, again, imply a form of psychologism. Presumably, questions (b) and (c) have a more intimate link. But it is not clear that there is but one possible justification of our logical norms, and that there is only one set of legitimate standards, either in deductive or in other sorts (e.g. probabilistic) of reasoning. Although an ecumenist pluralism seems to reign in many circles, the question: “Which logic is the right one?” still arises.

The received wisdom, therefore, is that the division of labour, which leaves to the psychologists the answer to (a), to the logicians an answer to (b), and to logicians and philosophers an answer to (c) is perfectly in order as it stands. But is this wisdom correct? Some *prima facie* doubts can be raised. If our logical norms had nothing to do with the way the actually reason—if logic were totally irrelevant to reasoning—how could they be prescriptive of our practice? How could we apply these norms? How could we teach logic? And if our reasoning practices had nothing to do with the way we evaluate reasoning, how could we simply describe the practices as a form of *reasoning*? In order to describe certain thought processes as reasoning and inference, we need at least some characterisation of what reasoning and inference are, and it is unclear that this can be done without logic. I believe that we cannot draw such a sharp line between our actual reasoning practice, which the object of psychology and the study of valid reasoning, which is the logician’s business. In this sense, I subscribe to a form of psychologism, although quite weak, and I have argued elsewhere for the view that, between the world of psychological processes (world 2 on Popper’s famous hierarchy of worlds, world 1 being that of physical objects) and the world of abstract thought (world 3, Frege’s “Third Realm”), there is room for an intermediary world 2, 5 situated midway between world 2 and world 3, between the muddy waters of psychology and the dry land of the logical norms (Engel [1989], [1996]). I shall not try here to defend this view in general, and I shall limit myself to deductive reasoning with propositions, leaving aside the large body of work and controversy about probabilistic reasoning.

I want to relate not only the questions (a)–(c) above, but also three kinds of concerns: those of the psychologists of reasoning, who are interested in whether we reason through formal rules or through other ways, in particular through the construction of mental models; those of logicians, who are interested in the relationship between a semantic and a proof-theoretic account of logical systems; and those of philosophers, who ask whether meaning is a matter of acceptance conditions (or assertion conditions) or of truth conditions. My aim is to sketch – in a quite programmatic way—what I take to be an integrated answer to this set of questions and concerns. This answer will be strongly influenced by Peacocke ([1988], [1993], [1992]).

This paper has three parts. In the first one I discuss some well known objections to the relevance of logic to reasoning, and suggest some ways in which this relevance can be assessed. In the second part, I propose a general framework for answering to the above questions and concerns. In the third part I deal, more briefly, with the relation between this view and familiar arguments about the rationality of human thinking and its justification through a “reflective equilibrium”, and indicate how they differ from these arguments.

I. *Is logic relevant to reasoning?*

It is a commonplace that everyday reasoning does not conform to the usual rules of classical logic, in so far as these are understood as descriptive of the practice of individuals. A large body of contemporary psychological work shows that people make systematic mistakes and fallacies. Philosophers and psychologists tend to agree on the fact that in so far as logic tells us which arguments or proofs are valid, it does not tell us anything about our reasoning as such, if by “reasoning” we mean *inferences from beliefs to other beliefs*. Among philosophers, for instance, Harman [1986] has argued that reasoning, as “reasoned change in view”, or as a process of belief revision, has nothing to do with logic. Harman takes the familiar rule of *modus ponens*: *P, and if P then Q, therefore Q*. As he points out, this principle says nothing about our beliefs. It does not say that if one *believes* that P, and also *believes* that if P then Q, then one can *believe* that Q. It does not even say that one should, given the first and the second belief, have the third. Sometimes, one should give up the belief that P, or the belief that *if P then Q*, instead. For instance, suppose you believe that if violets are red, then daffodils are, and that you also believe that violets are red. Should you believe that daffodils are red? No, you could drop the belief that violets are red, or the belief that if violets are red, then daffodils are. Harman considers also even more abstract logical principles, such as implication or non contradiction. Logic tells us that if a proposition implies logically another then we should accept this proposition. But the case just mentioned of the *modus ponens* shows that this need not be the case. The notorious problem of logical omniscience shows that one should not require, of a cognitive system, natural or artificial, that it believes or accepts everything that is implied logically by what it already believes or accepts, for it would have to believe or accept an infinity of propositions, which obviously finite minds such as us cannot do. And we should not clutter our minds with trivial implications. Similarly logic is supposed to prescribe us to avoid logical inconsistencies. But should we avoid any inconsistency when we encounter one? Suppose that someone is confronted to the Preface paradox: he writes a book of which he thinks that every sentence is true, but writes in his preface that he might well have made a mistake somewhere. He contradicts himself. But should he stop believing what his book says? This would be absurd. Harman concludes

that logic, as a theory of validity, has no special role to play in an account of reasoning. It is merely “a body of truths”, a science, like physics of chemistry (Harman [1984], 109). As such, it is not normative, it does not tell us anything about what we may or should *believe*.

Another familiar philosophical argument about the *normative* irrelevance of logic to reasoning and inference is the famous Lewis Carroll’s story of Achilles and the Tortoise (Carroll [1895]), which I have commented elsewhere (Engel [1998]). One of the main points of the story is that explicit knowledge of a valid truth of logic or of a valid rule of inference is insufficient to make the mind move, and there is always a gulf between this knowledge and our putting it into practice.

Psychologists concur with philosophers in the irrelevance of logic to everyday reasoning. Decades of work on Wason’s selection task seem to confirm this. And one of the most obvious lessons of work in this field seems to be that everyday reasoning is much more sensitive to the *content* of particular statements than formal logic is. Content can affect assessments of validity, and can impair good reasoning. In other cases, it has facilitating effects. *Content*, however, just seems to be another name for *belief content*. If a content is believable, then it tends to be believed, and its tends to be believed whether or not logic, or any sort of normative standards, tells us to do so or not to do so. In particular, even if an inference is invalid, but if its conclusion is highly believable, it tends to be accepted. For instance people accept easily the invalid but highly believable syllogism:

All of the Frenchmen were wine drinkers
Some of the wine drinkers are gourmets
∴ Some of the Frenchmen are gourmets

Studies in the “belief bias effect” strongly document this (Santamaria, Garcia-Madruga & Johnson Laird [1998]; Torrens, Thomson & Cramer [1999]). Another relevant field of study is work on “illusory inferences”. For instance Savary and Johnson-Laird [1998] have shown that the very processes of construction of mental models which, according to them, is the basic process of reasoning, leads people to consider as valid inferences which actually are invalid, such as:

If there is a king in the hand then there is an ace in the hand, or else
if there is queen in the hand then there is an ace in the hand.
There is king in the hand
∴ There is an ace in the hand

Other studies tends to show that features of the comprehension processes of the sentences involved in the tasks, or pragmatic features which are extraneous with respect to the logical forms of the inferential tasks, have more influence than the logical contents or logical forms of the inferences in the drawing of conclusion (see e.g. Evans [1994], Sperber, Cara and Girotto [1995]). The lesson of these studies seems clear: validity and believability fall apart, and it is one thing to assess the validity of

an inference, and another thing to assess its believability and the way people actually draw their conclusions. According to mental model theory, this is not surprising, since the assessment of validity often implies the construction of a large number of mental models, and of sophisticated techniques for reducing these constructions to the relevant cases, whereas the “natural” construction of models by untrained individuals allows only a limited amount of processing and of imagination. People tend to “satisfice” by taking up models which are readily available, neglecting others and failing to look further for more models which would confirm or infirm a conclusion.

Other psychologists, who hold different views about the actual processes of reasoning (see e.g. Gigerenzer [1995]) tend nevertheless to agree that it is wrong, when one assesses these processes, to look for the “logically correct answer” and to contrast it from what is *logically* right: it might be that subjects are right on other counts, and, in any case, there is no real question, when one deals with reasoning of who is “right” or “wrong” with respect to a given standard or norm. The question is rather, to see how people do actually deal with the tasks, and to separate the descriptive issue from the normative one. On such a view the norms of reasoning are not independent from our psychological capacities. There is not assessment of what is “rational” for an agent apart from the way he actually performs a given task. In other words, there are no other norms that those which emerge from the psychological capacities, studied in their social contexts. Norms have to “psychologized” (Gigerenzer [1998]). This comes very close to full-blown psychologism, and it is open to the current objections against such a view.¹

As I said above, although this separation of questions (a) from question (b) seems to be the natural division of labor, and that the freeing of the psychologist from the normative governance of the logician is in itself a good thing, I do not think that this division is correct. Psychologists like Gigerenzer, in their boldest moments, seem to speak as if we could completely separate off the issue of our normative standards of validity from the issue of our actual practice of reasoning. I suspect that mental model theorists can be tempted at times by this view. But it is not clear that, on most psychological theories of reasoning, reasoning has *nothing* to do at all with valid reasoning in the sense of an ideal standard of what *should* follow from what. On theories according to which people reason with formal rules of inference (e.g. Rips [1994]) indeed people use such valid rules. On mental model theory (henceforth MMT), people do not use such rules, but their reasoning has something to do with validity, at least on two counts:

a) in so far as validity is a *semantic* notion, MMT supposes that people use semantic notions: indeed they use the most basic one, *truth*, for

¹ Gigerenzer illustrates this mostly from probabilistic norms, but his view would probably be parallel with deductive reasoning.

they represent themselves the truth of the propositions they are given, and each model is a situation in which a proposition is represented as true

- b) the construction of models obeys the semantic principle of validity: that an inference is valid if and only if there is no way in which its premises could be true and its conclusion false. The construction of models is guided by the search for counterexamples. Johnson Laird and Byrne ([1991], 209) call this “the central core of rationality” In this respect MMT presupposes that people have at least an implicit understanding of validity, although the use of the semantic principle of validity does not imply that people have a tacit repertory of valid rules of inference, even less that they use such rules explicitly.

(Let us note, in passing, that what we may call the *semantic core* of reasoning according to MMT can be found also in the quasi-Tractarian theory of probability that it uses when it deals with probabilistic reasoning (see Girotto & Legrenzi [1998], 258).)

So even at the unreflective level of the representation of situations where the state of affairs described by a proposition obtains, MMT presupposes that subjects have at least an implicit grasp of two normative notions: truth, and validity. The use of these notions, however, occurs also at a reflective level. As Johnson-Laird and Byrne ([1991], 147) note:

Reasoners can *know* that they have made a valid deduction, and these meta-deductive intuitions prepare the way for the development of self-conscious methods of checking validity. without this higher-level, or meta-deductive, ability, human beings would not have invented logic, they could not make deductions about other people’s deductions, and they could not devise psychological theories of reasoning.

This tempers somewhat the radical claim that logic has nothing to do with reasoning. For we can understand this claim not as a statement to the effect that logic, as the reflective assessment of meta-deduction, is not situated at the same level as our untutored inferential intuitions, although it is necessarily connected with these intuitions: indeed we would not even understand the rules of inference and the principles of logical validity if they had absolutely no relationship with our intuitions about inference and validity. It follows from these remarks that we should distinguish two sorts of claims of irrelevance (and of relevance) of logic to actual reasoning:

- i) logic is irrelevant to reasoning because the usual logical laws or rules of inference (of classical logic) do not *describe* the actual processes of reasoning, and because subjects do not have an *explicit* representation of the laws
- ii) logic is irrelevant to reasoning because the usual logic laws and rules of inference (of classical logic) do not *prescribe* the way to reason, nor guide *implicitly* our actual reasoning.

From the few remarks above, I take (i) to be true, but (ii), although correct as it stands, is not obviously false. If by *prescription*, one means that a subject could have an explicit representation of logical rules, and “follow” them consciously, it is clear that most psychological findings show that people do not reason that way, at least when they are untrained. If MMT is correct, and if rule-based theories are incorrect, then it is not even true that we are implicitly guided by formal logical rules. Now from this it does not follow that usual logical rules and principles have no normative role to play in reasoning, for people at least tacitly use the principle of semantic validity in their construction of mental models, and must have *some* grasp of the rules that they use if they are to have any meta-deductive intuitions at all.

The distinction between (i) and (ii) seems to be close to Evans’s [1996] distinction between two sorts of rationality: rationality₁, the ability to think in the service of a goal, and rationality₂, the ability to reason according to a normative theory. But Evans considers that the two are strongly independent, for about the same reasons as those given above. His claim, therefore, looks very much like (i) as well as (ii) This is not my claim. My claim is that even at the level of untutored intuition, people have some grasp of *rationality*₂. They grasp a normative theory, even though they do not do this *explicitly* nor by consciously follow norms.

My aim is to suggest what grasp of a normative theory is involved here. There are two dimensions here. One concerns the content of the norms. The other one is the nature of the implicit understanding of these norms. The two are not independent from each other.

Let us try first to characterise the content of the norms involved. MM theorists have suggested that people have the grasp of two basic norms, truth and validity. Truth indeed is very weak and general, for according to what Johnson-Laird and his associates call “the principle of truth”, people tend to minimise the load of working memory by representing explicitly only what is true, not what is false. As they say: “The theory assumes that the mental footnotes [of the implicit models] about false cases are rapidly forgotten and all but the simplest of cases. There is a premium on truth in human reasoning. what is left out of models is explicit information about possibilities and literals that are *false*.” (Savary and Johnson-Laird [1998], ms 5). Now it seems odd to say that people possess the concept of truth in their manipulation of propositions if they do not take into account cases where the propositions which are given to them are false. Grasp of truth as a norm should go with grasp of the concept of falsity, and have an idea of the asymmetry between truth and falsity. Now MMT predicts that people *infer* the false cases from their knowledge of true cases. And in their application of the principle of semantic validity, they do look for counterexamples. For instance, even if they initially represent conditional statements through models where the representation of the falsity of the antecedent is only implicit, they

have a conception of what would falsify a conditional, namely a case where the antecedent is true and the consequent false.

According to MMT, there is a least this minimal understanding of the notion of truth and validity in reasoning. I have argued that it is enough to temper the claim that logical normativity has nothing to do with actual reasoning. Still, the amount of logicity involved is very weak. The question is whether we can ascribe to the subjects a better grasp of logical principles than this minimal understanding.

On the theory according to which people reason from their tacit knowledge of a basic set of formal rules of a mental logic, subjects do have more than this minimal understanding. They indeed are supposed to master and to use a certain amount of rules similar to those of natural deduction. But such views grant *too much* logicity to subjects. Against these views, MM theorists hold that only the minimal representations of models and the memory driven principles of construction are necessary. The amount of “logic” involved in the basic models for the propositional connectives postulated by MMT (see e.g. Johnson Laird and Byrne [1991], 51) is very small. MMT is a parsimonious theory of reasoning, and it owes its explanatory virtues to precisely this simplicity: when a theory explains more with less principles, the more powerful it is.

I want, however, to argue that more than this minimal amount of logicity is needed for reasoning, even though we do not need to suppose that people possess full blown formal rules. What is it, on MMT, to understand the meaning of a propositional connective, such as “and”, “or” or “if...then”? It is to understand a basic set of mental models for each connective, explicit and implicit. For instance

A and B	<i>implicit explicit</i>
A or B (inclusive) $A \vee B$	A B A B
	B \neg A B
	A B A B

It would be wrong to say that the meanings of connectives are *explicitly defined* by their canonical mental models, since the theory postulates that there are implicit definitions of these connectives. MM theorists insist that the construction of mental models is not equivalent to the use of rules. But even if we grant this, it is not obvious that subjects cannot form inferences from their understanding of the mental models associated with each connective. For instance a subject who is asked to consider a conjunction “A and B” and who represents the mental model A B above, is certainly able to assess the simple inference

A and B
therefore A

for A is indeed represented in his model. And the same for the inference

A and B
therefore B

Just in the same way, MMT shows how the subjects who have the mental models for the conditional “If A then B”

A B
...

can readily assess *modus ponens* forms of inference. As well known, the above inferences for “and” and “if” are called, in natural deduction systems, the elimination rules for each connective. But these kinds of inference are not just some among those which are easily constructed from the mental models. They are indeed partly constitutive of the *meaning* of the connectives. A subject who would not agree that “A” follows from “A and B” would simply show that he is not able to construct the appropriate mental model for the connective, and hence that he does not know the meaning of “and”. In this sense, the meaning of a connective is *constituted* by the role that it plays in reasoning. If one grants this, I do not see why MMT would be incompatible with the view that there are *canonical* kinds of inference attached to each connective. We could have an indirect confirmation of this if we considered a spurious connective, such as Prior’s [1960] *tonk*. Everyone knows the story about this odd connective. How could we tell it in the vocabulary of MMT? I suppose that it would go something like this. Imagine that I tell you that *tonk* is a connective for which when

A *tonk* B is true,
A is true

and for which when A *tonk* B is true, B is true.

On MMT subjects should construct the model

A B

You will be led to identify *tonk* with “and”. But now suppose that I tell you that it is also the case that when

A is true
A *tonk* B is true

(and similarly when B is true, A *tonk* B is true). You will then be unable to represent the truth of A *tonk* B, for if you form a model

A

you obviously cannot form the model

A B

from it. However, the former model is consistent with

A
B

But here you get lost, for the two models are incompatible. They do not represent the same state of affairs. In other terms, you cannot form

mental models from the rules given to you about this connective *tonk*, and this is why the connective has no role in one's conceptual economy. This could be also formulated in terms of classical truth tables and natural deduction rules. The introduction rule for *tonk* requires that "A *tonk* B" is true when A is true and B is false. Its elimination rule requires that in the same case "A *tonk* B" is false. Otherwise the rule will lead from truth to falsity. Hence there is no coherent truth table for *tonk*. I shall say more on this below, but for the moment I conclude from this that a subject who understands the meaning of a connective understands some basic rules of inference associated to the mental models that he is able to construct in inferences involving this connective, and that these rules are normative for his inferential practice. The norms attached to each connective are more substantial than the mere grasp of the concept of truth and of semantic validity.

The second dimension of normativity concerns, as I said, the nature of the implicit understanding of the norms involved. One of the reasons why logic, as a set of valid rules or principles, seems to be irrelevant to everyday reasoning, is that people do not have any explicit, reflective, knowledge of these rules; moreover, many psychological experiments show that even people trained in logic or in other normative theories commit systematic errors. This was also part of Lewis Carroll's point: explicit knowledge of a valid rule is of no help for reaching a given conclusion. This is why knowledge of the norms of reasoning must be implicit. But the alternative view, that people obey to blind compulsions, and are simply disposed to reach certain conclusions without any appreciation of what rules they are following, is also unattractive: if they are to reason at all, they must have some appreciation of the norms governing their performance. If the foregoing remarks are correct, even when a number of experiments seem to show that people jump to conclusions *because* of certain bias, memory constraints, and other causal factors, it does not follow that people fail to have *reasons* for inferring what they infer. So even if their understanding of inference rules is not explicit, it is not simply blind. (This remark has obvious ties with Wittgenstein's discussion of rule-following, but I cannot take it up here.²) An account is needed which would make room both for the implicit (and hence not completely conscious) grasp of the norms and for the fact that people have at least some knowledge of the reasons of their inferential activities.

II. *Logical constants and conceptual role*

Thus far I have only considered the relationship between questions (a) and (b) above, and argued that reasoning is not a completely norm free

² Wittgenstein argues *both* against the idea that we would have, in any rule-governed activity the explicit knowledge of the rule set up in front of our mind (for this would pose the threat of a regress, like in Carroll's story) *and* against the idea that following a rule is simply a matter of being disposed to act in certain ways.

See the interesting remarks of Brewer [1995].

activity. But I have completely left out question (c), the question of the *justification* of logical rules and principles. What justifies these rules? Can they even be justified? Are they based on more primitive principles, or are they themselves basic? Are they secure? Such questions (see e.g. Dummett [1973]) are situated at an even more higher level of abstraction than the question whether we are aware of the validity of principles of logic and really use some valid principles, for they deal with the nature of validity itself in this domain. Given that people do not use valid rules of reasoning such as those that logicians investigate in their everyday reasoning, why should they care about the very nature of validity? Of course they don't, unless they reach this highly sophisticated stage of meta-deductive inquiry which comes when logicians raise questions about the consistency, soundness and completeness of their rules. Nevertheless there is an actual link between the psychological problem of logical competence and the metatheoretical problem of the justification of deduction. It is this the following. To justify logical principles is to give an account of the meanings of the logical expressions which determine a given set of sentences and of inferences as "logical", i.e. of the "logical constants". There are basically two styles of justification of the principles of logic, or two ways of determining the meanings of the logical constants. One is proof-theoretic, and appeals to properties of the systems of rules, independently of such semantical notions as truth or validity. The other one is semantic, and appeals to such properties. The two are not incompatible, of course, but they can become incompatible if the truth or the validity of a given sentence (or set of sentences) is understood in terms of the *assertibility* of this sentence. Those writers who favour an anti-realist conception of truth, according to which there is no more to truth than assertibility, and who reject the idea that there can be true but unverifiable statements, will tend to favor a purely proof-theoretic treatment (for this debate, see e.g. Tennant [1987], Dummett [1991], Engel [1989]). Now there is a similar alternative between a proof-theoretic procedure of defining the logical constants and a semantic procedure in the psychology of reasoning, although this is not alternative about how these constants should be justified, but about whether they are used at the psychological level: it is the familiar alternative between "formal rules" accounts and "mental model" accounts. The former are clearly proof-theoretic, and the latter clearly semantic. But apart from this analogy, what is the connection between the two? The first alternative is a logical one, the second a psychological one. It does seem to me that there is an important connection between the two, and I want to argue that *the very reasons which justify a certain account of the meanings of logical constants are those which operate, at the psychological level, to rationalise the reasoning behaviour of subjects, and these reasons are semantical.*

Let us come back to Prior's *tonk*. Prior invented this spurious connective in order to show that the falsity of the thesis according to which log-

ical constants could be introduced by convention and through stipulated inference rules. In his equally famous reply to Prior, Belnap pointed out that what was wrong with *tonk* was not that its elimination and introduction rules were inconsistent, but that the requirement of conservative extension was not met: one should not introduce a new constant in a language if that constant does not allow derivations which were not allowed in the language before the introduction of this new constant. Now this conservative extension requirement is purely proof-theoretic, and it is emphasised by philosophers would favour an anti-realist view of logical justification (see Dummett [1991], 217–220, Tennant [1987]). But let us return to the remark about *tonk* above. The upshot of these remarks was that there is not coherent semantic assignment of truth values for *tonk* (and this why people cannot form mental models for *tonk*). So we can formulate what is wrong with *tonk* in semantic terms, rather than in proof-theoretic terms: there is no assignment which makes *tonk* inferences truth preserving. Indeed Peacocke ([1988], [1993]) has formulated a requirement of conservative extension in semantic terms. The proof-theoretic requirement is (for a set of rules S, and where L is a system of deduction for a language not containing the new logical constant):

S is deductively conservative over L iff for any sentences A1, ...An, B of L,

$A1, \dots, An \vdash_{L+S} B$ only if $A1, \dots, An \vdash_L B$

and the semantic requirement is:

S is semantically conservative over L iff for any sentences A1, ...An, B of L,

$A1, \dots, An \models_{L+S} B$ only if $A1, \dots, An \models_L B$ (where ‘ \models ’ is the relation of semantic consequence)

Peacocke then argues that conservative extension is a requirement which is available for someone who favours a realist view of truth and a semantic justification of deduction. I shall not develop here his argument (see Engel [1989], ch. XII). The important thing for us here is that we can translate this framework for a theory of the justification of deduction into a framework for a theory of the meaning and of the understanding of logical constants.

This appropriate framework is that of a semantic in terms of “conceptual roles”. A conceptual role semantics is in general a theory of meaning according to which to explain the meaning of an expression is to specify its role in reasoning and inferences. There are a number of versions of this view (see for instance Harman [1982], Block [1986]) but the one which interest us here is this: the conceptual role of an expression is constituted by what leads a subject to accept certain kinds of sentences (the acceptance conditions of a sentence), and to infer certain kinds of sentences from others. In this sense it is close to a “use” theory of meaning, and even to a verificationist view. In the case of the meanings of the logical constants, a pure conceptual role theory of this kind amount to the thesis that their role is entirely given by their elimi-

nation and introduction rules, without any appeal to semantic notions such as truth or validity. But we have just seen that such a pure theory can be formulated in semantic terms. Indeed, if we generalize from the case of *tonk*, it can be argued that the acceptance conditions (or the assertion conditions) of a given connective *determine* its truth conditions. In other terms, a conceptual genuinely determines a meaning for an expression only if there is a semantic value of the appropriate category for this expression which makes the transitions mentioned in the conceptual role always truth-preserving. (In Peacocke’s [1992] terminology, this amount to saying that there is a “determination theory” for a given concept.)

On Peacocke’s framework, which I follow here, understanding a logical constant is understanding these principles for it which makes them truth-preserving. Peacocke holds also that this understanding is constituted by finding the introduction and elimination rules “primitively obvious” For instance, in the case of negation, Peacocke says that

What is primitively obvious to anyone who understands negation is just that A is incompatible with A. Unless the ordinary user of negation appreciates that A and $\neg A$ cannot be both true, then he does not understand \neg . ([1988], 163)

Peacocke further hypothesizes that in order to find this “obvious”, the subject must have an “implicit conception” of negation. There is no incoherence in the idea that one can have an implicit or tacit knowledge of a normative rule. Only the thought that the alternative must be between conscious awareness of the rules and brute compulsion can lead us to think the contrary. Thus we have three conditions:

- (i) the rules for the basic connectives are “obvious”
- (ii) they are normative, for they flow from their conceptual role
- (iii) this conceptual determines truth conditions and validates the connective introduction and elimination rule.

Now (i) is an answer to question (a) above, (ii) is an answer to question (b), and (iii) is an answer to question (c). There is a parallelism with the connectives that we find primitively obvious, the fact that they give us reasons to draw a certain number of inferences, and the fact that the inference rules of the connectives are semantically truth-preserving, hence justified.

I anticipate the reaction: isn’t it too good to be true? Have I been cheating? It might be objected also that in insisting, with Peacocke, on the idea that the meanings of the logical connectives are given by their basic inference rules and their conceptual roles, my view just amounts to a revamping of the traditional conception of a mental logic of formal rules. To the first question, I shall answer that my view implies that there indeed a connection between the validity of a logical rule (the fact that it is truth preserving) and the impression of validity that people have. Otherwise, it seems to me, we could not recognise some rules, such as *modus ponens*, or the elimination and introduction rules for

“and” as normative: we could not move from our implicit understanding to an explicit understanding of the connectives. But of course, as logicians and psychologists alike insist, it is not because a rule *seems to us* to be valid that it *is* valid, and it is not because a rule *is* valid that it will *seem* to us to be so. If the former were the case, psychologism would be true, and I agree that it is not, as an attempt to derive validity from impression of validity. If the latter were the case, we would not observe such a gap between logic and reasoning competence. But the relationship between impressions of validity and validity has nevertheless to be close enough, if the meaning of logical words is supposed to be normative. To the second question, I answer that my view does not amount to a mental logic view. Nothing in what I have said implies that subjects have a full mastery of all the elimination and introduction rules for the connectives that natural deduction logic postulates. But they must have at least a partial grasp of it, and must be able to recover some of these rules from (e.g. *modus ponens*) their intuitive understanding. Furthermore nothing in what I have said implies that there is some algorithm which could allow subject to recover the truth conditions (and hence the mental models by which we construct them, on MMT theory) from their inferential rules and their conceptual roles. It does not, therefore, seem to me that there is an incompatibility between the present version of conceptual role semantics and mental model theory.

III. *Rationality and reflective equilibrium*

I have claimed that there is a form of harmony between our reflective understanding of logical rules and the norms of logic, in spite of the fact that people do not reason according to logic rules. This position may seem to be close to the position of those writers who have defended a similar pre-established harmony under the name of “reflective equilibrium” (Cohen [1998]). According to this conception there should not be any real discrepancy between the normative principles of reasoning (logic) and human reasoning competence since these come from our intuitions about what constitutes good reasoning, which in turn comes from our reasoning competence, since our normative principles are the product of a series of weighted comparisons between these and our intuitions, until an equilibrium is reached. Reflective equilibrium thus has been used as an argument for the claim that there is an *a priori* presumption of rationality, a sort of virtuous circularity between our intuitions and our normative standards. Other arguments for this claim are the use of various principles of charity in interpretation which are supposed to give us an *a priori* warrant that subjects cannot but be rational (for a study, see Engel [1993])

Reflective equilibrium arguments and similar arguments about the necessary harmony between logical competence and the normative rules of logic have been strongly criticized (see Stein [1996]). On the one hand, what warrant do we have that the intuitions which are supposed

to serve as an input in the reflective equilibrium method are rational? Indeed many of our intuitions are wrong. By what miracle should we in the end come up with an harmony between these intuitions and the normative principles? On the other hand, if the reflective equilibrium method just supposes that in the end, after a series of appropriate adjustments we reach this harmony, then it seems to simply beg the question.

The present argument, however, does not depend upon the notion of a reflective equilibrium, nor upon such arguments as those which use the principle of charity in interpretation. It does not depend on the first because, unlike the reflective equilibrium method, a conceptual role theory of meaning is not strongly holistic. It is holistic in this sense that it supposes that the meaning of an expression is determined by its inferential connections with other expressions, but it is not *strongly* holistic in the sense that it would have to suppose that *all* the inferential links of an expression matter in fixing its conceptual role. This thesis is often called “meaning holism”, and the implication just mentioned is not found, to say the least, very attractive by many.³ The thesis defended here, with respect to logical constants, is that a *delimited* set of inferential roles is characteristic of a logical constant, and serves to define it implicitly. Hence this thesis grants, against strong holists such as Quine, that there is a legitimate distinction between a set of meaning giving principles which are *analytic* and other meaning link which are synthetic. This is not the place to argue for this here, but it seems to me to be an implication of this view. For just the same reason, the present view does not depend upon a method of radical interpretation based on the principle of charity, and hence it does not allow any argument to the effect that there is a general presumption of rationality.

Conclusion

I have been far too sketchy. But let me try to summarise nevertheless what I have attempted to propose. First, in spite of the fact that logic, as a normative discipline made up of valid principles and rules, is not relevant to everyday reasoning, it is not true that it is totally irrelevant. Even psychologists who, like the MM theorists, claim that reasoning goes by construction of models and various bias and not by formal rules, admit that we could not have meta-deductive intuitions if everyday reasoning did not at least allow people to have a grasp of the normative concepts of truth and semantic validity. Second, it seems to me that we can go further, and argue that people have an implicit understanding of some normative logical principles, at least a subset of the inferential rules for the logical connectives, and that this understanding is based on a semantic principle, according to which a logical connective is legiti-

³ See Stein’s (1996) comments p.124-127 on the relationship between the principle of charity and holism.

mate if and only if it is truth preserving. There is thus a link between a realist view of truth (truth is not assertibility), a semantic justification of logic, and our implicit understanding of normative principles. It seems to me that there is nothing here with which semantically inspired researchers on reasoning could disagree. The gap between logic and the psychology of reasoning is not, on my view, as large as it is often claimed to be.^{4*}

References

- Belnap, N. [1961–2] “Tonk, Plonk and Plink”, *Analysis*, XXXII, 130–134.
- Block, N. [1986], “Advertisement for a Semantics for Psychology”, *Midwest Studies in Philosophy*, X.
- Brewer, B. [1995], “Compulsion by Reason”, *Proceedings of the Aristotelian Society*, sup. vol. LXIX, 237–254.
- Carroll, L. [1895], “What the tortoise said to Achilles”, *Mind*, 4.
- Castellani, F. & Montecucco, L. [1998], a cura di, *Normatività logica e ragionamento di senso comune* (Bologna: Il Mulino).
- Cohen, L. J. [1981], “Can Human Irrationality be Experimentally Demonstrated?”, *Behavioral and Brain Sciences*, 4, 317–331.
- Dummett, M. [1973], “The Justification of Deduction”, in *Truth and Other Enigmas* (Cambridge, Mass.: Harvard University Press).
- [1991], *The Logical Basis of Metaphysics* (Cambridge, Mass.: Harvard University Press).
- Engel, P. [1989], *La norme du vrai* (Paris: Gallimard), revised English tr. *The Norm of Truth* (Toronto: Harvester Wheatsheaf and University of Toronto Press, 1991).
- [1993], “Logique, raisonnement et rationalité”, in Houdé, O. et Miéville, D., *Pensée Logico-mathématique, nouveaux objets interdisciplinaires* (Paris: P.U.F.), 205–228.
- [1996], *Philosophie et psychologie* (Paris: Gallimard, Folio), Italian translation *Filosofia e psicologia* (Torino: Einaudi, 1999).
- [1998], “La logique peut-elle mouvoir l’esprit”, *Dialogue*, XXXVII, 35–54.

^{4*} This paper was read at the conference “Reasoning: the logical and the psychological perspectives”, at the University of Padova, in May 1999. But it is actually a descendent of a paper, on Lewis Carroll’s paradox, read two years earlier in the same place at the conference “Modelli mentali” in honor of Philip Johnson Laird, which was published in a French version as Engel 1998. In both occasions, Philip Johnson Laird’s comments have been enormously useful, and my debt to him goes further. He is the one who put me to read more of the psychological literature and his open mindedness to philosophical issues is unique. I thank also for their remarks Marco Santambrogio, my commentator, Vittorio Girotto, Paolo Legrenzi, Daniele Giarretta, Giorgio Samblin, Paolo Leonardi, and Carlo Filotico; for their hospitality on both occasions Alberto Mazocco, Paolo Flores d’Arcais, and Daniele Giarretta, and Luisa Montecucco for all her encouragements. An Italian version of the text has been published as “Logica, ragionamento e costanti logiche”, ch. 3 de P. Cherubini, P. Giarretta et A. Mazocco (eds.), *Ragionamento, psicologia e logica* (Firenze: Giunti, 2001), 108–127.

- Evans, J. St. B. T. [1993], “Bias and Rationality” in J. Manktelow & D. Over, *Rationality* (London: Routledge).
- [1994], “Relevance and reasoning”, in S. Newstead & J. St. B. Evans, eds., *Current Directions in Thinking and Reasoning* (Erlbaum: Hillsdale).
- Gigerenzer, G. [1995], “The taming of content: some thoughts about domains and modules”, *Thinking and Reasoning*, vol. 1, 289–400.
- [1998], “Psychological challenges for normative models”, in D. Gabbay and P. Smets, eds., *Handbook of Defeasible Reasoning and Uncertainty Managements Systems*, vol. 1, 441–467.
- Girotto, V. [1994], *Il ragionamento* (Bologna: Il Mulino).
- & P. Legrenzi [1998], “Logica, probabilità e ragionamento ingenuo”, in Castellani & Montecucco [1998], 241–260.
- Harman, G. [1982], “Conceptual Role Semantics”, *Notre Dame Journal of Formal Logic*, 23, 242–56.
- [1986], *Change in View* (Cambridge, Mass.: MIT Press).
- Johnson-Laird, P. & R. Byrne [1991], *Deduction* (Hillsdale, NJ: Lawrence Erlbaum).
- & F. Savary [1999], “Illusory inferences: a novel class of erroneous deductions”, *Cognition*, 71, 191–229.
- Montecucco, L. [1998], “Normatività e descrittività nello studio del ragionamento”, in Luisa Montecucco, *Contesti filosofici della scienza* (Brescia: La scuola).
- Peacocke, C. [1988], “Understanding Logical Constants”, *Proceedings of the British Academy*, 73, 153–200.
- [1992], *A Study of concepts* (Cambridge, Mass.: MIT Press).
- [1993], “Proof and Truth”, in J. Haldane and C. Wright, *Realism and Reason* (Oxford: Oxford University Press).
- Prior, A. [1960], “The Runabout inference ticket”, *Analysis*, XXI, 38–39.
- Rips, L. [1993], *The Psychology of Proof* (Cambridge, Mass.: MIT Press).
- Santamaria, C., J. Garcia-Madruga, & P. Johnson-Laird [1998], “Reasoning from double conditionals: the effects of logical Structure and Believability”, *Thinking and Reasoning*, 4, 97–122.
- Sperber, D., F. Cara & V. Girotto [1995], “Relevance Theory Explains the Selection Task”, *Cognition*, 31, 31–95.
- Stein, E. [1996], *Without Food Reason: The Rationality Debate in Philosophy and Cognitive Science* (Oxford: Oxford University Press).