Logical reasoning

Bart Geurts
Department of philosophy
University of Nijmegen
bart.geurts@phil.kun.nl

Abstract
In the psychological literature on reasoning it has always been assumed that if there is such a thing as mental logic, it must be a set of inference rules. This proof-theoretic conception of mental logic is compatible with but doesn't do justice to what, according to most logicians, logic is about. Thus, the ongoing debate over mental logic is based on a too narrow notion of logic. Adopting the broader perspective suggested by the standard (Tarskian) view on logic helps to clarify the debate and also shows that the case for mental logic is much stronger than its critics would have us believe.

1 Introduction

This paper is about the extent to which everyday reasoning may be called a logical process. I am not the first to broach this issue. The very first psychological theories of reasoning properly so-called all maintained that reasoning is a matter of applying inference rules.[1] But the sway once held by logic-based approaches has waned, if only because in the meantime so many alternative frameworks and programs have emerged. The main non-logical frameworks are Johnson Laird et al.'s theory of mental models and Chater and Oaksford's probabilistic approach.[2] In the rubric of alternative programs

[1] Pioneers of the rule-based approach to reasoning are Osherson (1977), Johnson-Laird (1975), Rips (1983), and Braine and his associates (Braine and Rumain 1983, Braine, Reiser, and Rumain 1984). To date, the most systematic and comprehensive proposal in this tradition is Rips's (1994).

[2] Chater and Oaksford's probabilistic approach is set out in Chater and Oaksford (1999) and Oaksford and Chater (2001). The literature on mental models is immense. Book-length introductions are provided by Johnson-Laird (1983) and by Johnson-Laird and Byrne (1991); a recent attempt at formulating a formally

I would place various pragmatic and evolutionary approaches, which cover only a narrow range of phenomena, and confine their attention to isolated instances of reasoning, such as the Wason task.[3]

The gradual decline of the logicist paradigm was accompanied by a controversy over the relation between logic and reasoning, with opinions lining up along a gradient ranging from the view that nothing could be less relevant to psychology than logic, on the one hand, to the view that logic is central not only to reasoning but to cognition at large, on the other. Recently this discussion has abated somewhat, but there has been no clearly definable outcome one way or another. However, and this is the central tenet of the present paper, the discussion about mental logic has always been framed the wrong way, because it was predicated on an overly narrow conception of what logic is: somewhere along the way logic got conflated with one of its branches, to wit proof theory. This is not surprising, since virtually all avowedly logic-based theories assume that reasoning is a process of applying inference rules, and typically take their lead from proof-theoretical systems such as natural deduction.[4] As a consequence, it has been taken for granted by all parties to the debate that what is at stake is whether or not inference rules are implicated in natural reasoning, the tacit assumption being that logic is about systems of inference rules. However, there is more to logic than inference rules, and many logicians and philosophers of logic would say that systems of inference, though certainly important, are not even essential to logic.

Ever since its inception in ancient Greece, it has been recognised that logic is about validity; logic is the science of valid arguments. Proof theory is a way of approaching validity, but it is not the only way. So if the question is whether, or to what extent, everyday reasoning is a logical process, the most obvious starting point is the concept of validity. Adopting this course will not only clarify the relationship between

---

explicit version of the theory is Bara, Bucciarelli, and Lombardo (2001). My characterisation of mental-model theory as "non-logical" is to be taken with a grain of salt. I believe that there are strong affinities between the theory of mental models and logical theories, and that in certain important respects mental-model theory represents a logical approach to reasoning, too. However, an in-depth discussion of this issue would fall outside the scope of this paper.

[3] See e.g. Cheng and Holyoak (1985, 1989), Cosmides (1989).

[4] See any of the references given in footnote 1.

logic and psychology, it will also throw new light on various issues in the mental-logic debate.

Before we get down to business there are two things that I would like to get out of the way. First, this paper is not a plea for a logicist approach to cognitive pyschology at large. My concern is with reasoning, and I will defend the position that logic is an essential ingredient in human reasoning, but that is as far as I will go on this occasion. Secondly, it is not part of my brief to argue that reasoning (or any particular kind of reasoning) is nothing but logical. My impression is that most of the researchers currently working in the field would say that reasoning is a hybrid affair, and that is my position, too. Therefore, what I will claim is not that human reasoning is logical, but, rather more modestly, that logic plays a substantial role in certain kinds of reasoning. But my principal objective is to clarify the debate on mental logic.

2 Logical validity

Logical validity, or validity for short, is a modal relation. If an argument is valid its conclusion follows from its premisses by necessity. The converse does not hold: not every necessary consequence is a logical one. Consider the following argument: “Fido is a poodle; therefore Fido is a dog.” Given that “poodle” is a hyponym of “dog”, the conclusion follows necessarily from the premiss; but it is not a logical consequence. The general consensus among logicians has always been that validity is contingent solely on the *logical form* of an argument: the logical constants it contains and the way they are combined; the meanings of non-logical terms like “poodle” or “dog” are not to the point. Hence, the pre-theoretical notion of logical validity comes down to this:

> An argument “p_1, ..., p_n ; therefore q” is logically valid if and only if it is impossible for p_1, ..., p_n to be true and q to be false at the same time, regardless how the non-logical expressions in p_1, ..., p_n and q are interpreted.

This is what I take to be the received view on logical validity. It goes back at least as far as Aristotle, but since Tarski (1936) was the first to clearly articulate it and insist that it be recognised as the key concept in logic, I will henceforth speak of the Tarskian view on logical validity and its place in logic. Later on I will use the adjective also to refer to the perspective on the psychology of reasoning that I will be arguing for, which is a

natural extension of the Tarskian view on logic. Tarski's analysis of logical validity has recently been the focus of a debate ignited by Etchemendy's (1990) attack on Tarski's 1936 paper,[5] so I should like to emphasise that I am concerned here only with Tarski's observations on the pre-theoretical concepts of logic and logical validity, and these have not been in the line of fire.

The two main ingredients in the concept of validity are necessity and formality. The logical form of an argument is determined by the logical constants it contains and the way they are put together. Strictly speaking, however, it is misleading to refer to *the* logical form of an argument. At the end of the day, the distinction between logical and non-logical expressions is a pragmatic one. Logicians choose their logical vocabulary with a view to boosting the generality of their systems, but there is nothing inherently logical about such words as "and", "all", or "not"; there is nothing hewn in stone about which terms are logical.

To illustrate this point, consider the branch of logic that is concerned with modal notions like necessity, possibility, obligation, and so on. What is it that justifies the enterprise of modal logic? Whereas it may be argued that conjunction, negation, and quantification are essential to science and human knowledge generally, the same cannot be said of modal notions. Words like "and" or "not" are rife in everyday and scientific discourse alike; modal words such as "necessarily" and "possibly" are not. So what is that these words have in common with "and" and "or" that makes modal logic a worthwhile undertaking? The answer is simple: it is sufficiently rewarding to view modal words as logical expressions, as elements of logical form. There is no principled reason why we shouldn't apply logical analysis to ordinary content words, like "umbrella" or "francophone". It is just that the logic of these terms is bound to be trivial: the set of logical consequences that we would gain by introducing a logical constant for "umbrella" is, if not empty, entirely devoid of interest. The logical constants of classical and modal logic are just more interesting, from a logical point of view, than "umbrella" is. Whatever logical form is, it is not an inherent property of arguments or their parts.[6]

A further observation regarding logical form is the following. Since validity holds by virtue of the logical form of an argument, it is tempting (and people succumbed to

---

[5] See e.g. Ray (1996), Sher (1996), Hanson (1997).

[6] For further discussion of his theme, see Tarski (1936), Hanson (1997).

the temptation) to conclude that validity has nothing to do with content. This is a mistake. Granted, if q follows logically from p_1, ..., p_n, then it doesn't matter how the non-logical parts of p_1, ..., p_n and q are interpreted. But of course the interpretation of the logical parts is absolutely essential; whether or not (say) "not not p" logically implies p depends entirely on what we take "not" to mean. Validity is a semantic notion *par excellence*, and if we wanted to highlight that fact, we could define validity as follows:

> An argument "p_1, ..., p_n ; therefore q" is logically valid if and only if the interpretation of the logical form of p_1, ..., p_n and q suffices to guarantee that q cannot possibly be false if p_1, ..., p_n are true.

This is equivalent to the definition given earlier, but it has the advantage of being more explicit about the importance of interpretation.

Putting together the foregoing observations about logical form, we must conclude that the subject matter of logic is anything but cut-and-dried. For one thing, pre-theoretical intuitions about validity, which is what logicians have to go on, are based on meaning, which is a slippery notion if anything is. For another thing, there is no principled criterion for drawing the boundary between logical and non-logical arguments. This is not to say that logic is a fuzzier business than is generally believed. It is merely to say that there is no such thing as *the* logic of X, no matter what we substitute for X.

3 When is reasoning logical?

All participants in the ongoing debate about mental logic appear to agree about what is at stake: it is whether people do or do not use inference rules. I grant that this question is important, but I submit that it is not the main one. As we have just seen, logic is about validity. Systems of inference rules are one way of studying validity, but there are other ways, too. So the main question, I submit, is whether people have a notion of logical validity and use it in everyday reasoning.

To show that these are really two distinct issues, and that the distinction matters, let us consider the humble logic of "and". Like its counterparts in other languages, the English word "and" is used in a variety of ways, but it has long been realised that some

of these ways license inferences that are logically valid. For example, it is pretty evident that p must be true if "p and q" is true, regardless what p and q stand for. This is where the logician's business begins. The standard logical treatment of "and" will be familiar, I trust. To begin with, our pre-theoretical intuitions regarding "and" are made explicit by treating conjunction as a truth function: "p and q" is true if p and q are both true, and false otherwise. This definition regiments the intuitive meaning of "and" while at the same time yielding a formally explicit version of the pre-theoretical concept of validity (in so far as it concerns conjunctive statements). Once we have a clear notion of the kind of inferences "and" licenses, we can ask ourselves how to actually obtain those inferences. For this purpose, we can devise a system of inference rules, like the following:

$$\frac{p \text{ and } q}{p} \qquad \frac{p \text{ and } q}{q} \qquad \frac{p \quad q}{p \text{ and } q}$$

These rules are designed to conform to the interpretation of "and", and they do: they allow us to derive sentences p and q whenever "p and q" is given, and to derive "p and q" whenever p is given and q is given. Crucially, however, we do not need these rules, or any rules, for making such inferences. For instead of using inference rules we can use the truth table associated with "and" (that is to say, the interpretation we have settled on) for the same purpose. Using truth tables for deriving conclusions from a given set of premisses may not always be as efficient as using a system of rules, but properly done it will deliver the same results.

What I have just said applied to logicians' logic, but it carries over to mental logic. Consider two persons who use different methods for reasoning with conjunctive statements: one uses inference rules like the ones above while the other uses his mental representation of the meaning of "and" (assume for concreteness that the second person has a mental representation of the truth table for conjunction). Both methods are logical: they rely on logical form for making inferences. The difference is just that one method uses inference rules while the other does not.

It may be objected at this juncture that, for all we know, the distinction between semantic representations and inference rules is a specious one. It might turn out, after all, that in point of fact the mental representation of the meaning of "and" *is* a collection

of inference rules.[7] This is correct, but it is not directly relevant to my present concerns. What I want to show is that using inference rules is only one way of reasoning logically, and for this purpose it suffices that the semantic representation of "and" might not be a set of inference rules. (Besides, I believe that the plausibility of the proof-theoretic view on meaning quickly starts to diminish once we go beyond the stock-in-trade inventory of logical constants; but that is another story.)

My proposal is that, before asking whether people use inference rules, we first ask the more general question if, and to what extent, they employ the notion of logical validity. The main reason for framing the issue thus is that it focuses on what is generally taken to be the essence of logic, and avoids the narrow proof-theoretical view of logic as inference rules. There are more advantages though. As the example of "and" shows, the Tarskian perspective on reasoning that I advocate may afford a more parsimonious treatment of inference than the logic-as-rules view; it does not require that we posit a separate mental faculty for explaining logical reasoning; and it forges a natural link between the interpretation of "logical words" such as "and", "all", and "not", on the one hand, and the role they play in reasoning, on the other. All of this will be discussed in greater detail in the following.

It bears emphasising, perhaps, that the foregoing observations are not directed against rule-based accounts of reasoning. Rule systems are one way of implementing the Tarskian view on reasoning, and I am persuaded that logical reasoning sometimes involves inference rules. But, as we will see presently, most of the available evidence for mental logic is not evidence for inference rules.

It follows from the general description of logical reasoning given above that, broadly speaking, there are three kinds of evidence which bear upon the question if, or to what extent, people may be said to reason logically. First, since validity is the key notion in logic, the most obvious question to ask is whether people are attuned to the distinction between valid and invalid arguments. Secondly and thirdly, since necessity and logical form are the principal ingredients in the concept of logical validity, we should ask whether these notions play a role in reasoning.

To begin with, does performance in reasoning experiments show any effects of the distinction between valid and invalid arguments? This question can be answered with a resounding yes. For example, logical validity is a potent factor in determining

[7] As is maintained by Braine and O'Brien (1991), for instance.

performance on syllogistic tasks. This is nowhere clearer than in one of Rips's (1994) studies, according to which the 22 most frequently endorsed syllogisms are all valid. Since, of the full set of 256 syllogisms, only 24 are valid altogether, this is pretty strong evidence for the significance of logical validity, and it is corroborated by virtually all studies on syllogistic reasoning.[8] The same can be said of studies that have used propositional reasoning tasks. For example, people rarely make mistakes with modus ponens or the simple inferences involving conjunction that we discussed earlier. Generally speaking, valid arguments prove to be easier than invalid ones. To be sure, people make lots of mistakes; they aren't logic machines. But nor are they indifferent to the distinction between valid and invalid arguments; which is what one should expect if they reasoned logically.

The second question is whether people are responsive to the fact that valid arguments are nececessarily correct.To the best of my knowledge, there has been only one study that addressed this question, and then only obliquely. In an experiment on syllogistic reasoning, Dickstein (1981) distinguished between two levels of instructions: one that "did not contain any reference to the principle of necessity in drawing a conclusion from the premises" (p. 230), and a second one that did. Dickstein's main finding was "... that clarification of the role of necessity in deductive inference significantly improves performance." (p. 231) This is what one should expect if people used the notion of logical validity, though it is a lonely piece of evidence, and open to alternative construals.

The third question is the most important one. By definition, someone who reasons logically relies on logical form. From this it follows that to the extent that people reason logically:

- we might find priming or training effects: when two tasks are of the same or similar logical form, performance on one could influence performance on the other;
- verbal protocols might supply evidence that people abstract away from non-logical content;
- we might find logical errors;
- we shouldn't find content effects (such as differences between familiar and unfamiliar or between abstract and concrete materials);

[8] The evidence is discussed at some length in Geurts (2003).

All of these diagnostics have played some role in the discussion about reasoning, though the last one has received more attention than the others taken together. They were also discussed at length in an article by Smith, Langston, and Nisbett (1992), as part of their defence of the claim that people use inference rules. However, none of these diagnostics has any direct bearing on that issue; what they test for is whether reasoning involves logical form. Smith et al. recognise this. They divide the question whether people use inference rules into two sub-questions. In the terminology of the present paper, their questions are: whether logical form plays a role in reasoning, and whether people use rules that exploit logical form. The bulk of Smith et al.'s paper addresses the first question, and they note that at the time of writing, "there is comparatively little empirical data on reasoning that can be brought to bear on the rule-application issue," (p. 5) which is still true, as far as I can tell. If this is so, it will be difficult to make a good case for inference rules, which is what Smith et al. aimed for. But it is still possible to achieve the more general objective of establishing that people use logical form, hence reason logically.

In the following I will confine my attention to the last two items of the above list: logical errors and content effects. I will not say anything further about priming/training effects or verbal protocols, simply because there is too little in the way of systematic data to make discussion worthwhile.

4 Logical errors

In everyday life, the epithet "logical" connotes such qualities as efficiency, objectivity, and irrefutability. Perhaps these connotations go some way to explain the popular complaint that, generally speaking, mental-logic theories are poor at accounting for reasoning errors.[9] As a general objection (as which it is sometimes intended) this is entirely without justification. Even if it is the case that for certain reasoning errors non-logical theories provide better accounts than what mental-logic theories currently have to offer, the notion that the latter have a general problem with errors has never been substantiated.

[9] E.g. Cheng and Holyoak (1985), Johnson-Laird (1983, 1997), Johnson-Laird and Byrne (1991), Bara, Bucciarelli, and Lombardo (2001).

Some advocates of non-logical approaches to reasoning believe that people are quite poor at reasoning, and that it is therefore a priori unlikely that there is a significant role for mental logic to play.[10] This is a non sequitur, and it is based on a dubious premiss, to boot. It is true, of course, that several decades of experimental research have shown that people aren't good at many reasoning tasks. However, it simply doesn't follow from this that in general people are poor reasoners. For one thing, one shouldn't forget that, from a methodological point of view, errors are often more interesting than normative behaviour. Experimental researchers have always *wanted* to find reasoning errors, so it doesn't come as a surprise that they have found them. For another thing, the positive side of the balance bears some weight, too: there are plenty of reasoning tasks that people are quite good at. Furthermore, even if it were true that people are poor reasoners, that wouldn't even begin to show that they didn't employ a mental logic.

For someone who takes the view that reasoning is a matter of applying inference rules, there are two strategies for dealing with errors. To begin with, it can always be claimed that an error has nothing to with reasoning per se, but is due, for instance, to problems of interpretation. In many cases, this course is entirely appropriate, and it has been adopted by logical and non-logical theories alike, though the former are more often chastised for doing so.[11] But if it is agreed that the error is due to faulty reasoning, the preferred line of attack is likely to be as follows:

> Errors arise naturally with respect to [logical inference rules], since individuals' processing limits and biases, competition from other nondeductive strategies, uncertainties about the correctness of a rule, and assumptions about the appropriateness of a rule can all interfere with the way people use them. (Rips 1994: 379)

This is the well-nigh inevitable consequence of adopting the view that logical reasoning involves the use of inference rules; critics of rule-based accounts tend to dislike it. Johnson-Laird (1997), for example, commenting on the passage just cited, states that it "... is odd that rules that are supposed to be intuitively sound can get messed around in all of these ways." This is not a knock-down argument, to be sure. Rips's position is

---

[10] E.g. Cheng and Holyoak (1985), Cosmides (1989).

[11] E.g. Evans, Newstead, and Byrne (1993), Johnson-Laird (1997).

perfectly consistent in itself, and not without precedent in other departments of psychology. It is widely accepted, for instance, that people use grammar rules, which apparently "get messed around" in all sorts of ways, too, and it might be held that there is no principled difference between grammar rules and inference rules.[12] Nevertheless, I sympathise with Johnson-Laird's complaint: it *is* peculiar that we should routinely use inference rules that we don't trust or are uncertain about, especially when some of these rules are of a quite elementary sort. But I also believe that within a Tarskian perspective on logical reasoning we can save the essence of Rips's position. Let me explain this with the help of an example.

It is a familiar fact that, in some reasoning tasks, people tend to derive "Q B are A" from "Q A are B", even when the inference is invalid because the quantifier Q is non-symmetric. In an experimental task with single-premiss arguments, about one third of the participants will claim, incorrectly, that "all B are A" follows from "all A are B", and as many as two thirds will endorse the conversion of "some A are not B" (Newstead and Griggs 1983). The idea that illicit conversion accounts for a major share of the errors in syllogistic reasoning goes back to Wilkins (1928), and is confirmed by the observation that the most frequent errors are all attributable to illicit conversion.[13] So illicit conversion is a fact of life. What are we to make of it on a rule-based approach? Following Rips's advice, we could hypothesise that there is a general rule of conversion, which is occasionally misapplied. But this raises worries of the kind voiced by Johnson-Laird. Why should there be a general rule of conversion? Why should people sometimes be mistaken about the applicability of this rule? And so on.

Taking a Tarskian stance on logical reasoning, we can adopt a different, and perhaps more plausible, view on the matter. Suppose we define the truth-conditional meanings of "some" and "all" along the following lines:

> "Some A are B" is true if and only if there is at least one A that is also a B.
> "All A are B" is true if and only if the set of A's is contained in the set of B's.

---

[12] Osherson (1977) is quite explicit about this.

[13] For further discussion of illicit conversion, see Newstead (1989), Evans, Newstead, and Byrne (1993), Geurts (2003).

Suppose, furthermore, that a given person (Mrs Conclusion, to give her a name) maintains a mental representation of these meanings. So Mrs Conclusion's semantic representation of "some" licenses conversion while her semantic representation of "all" does not. Now we present Mrs C with an experimental task: she is to decide whether or not "Q A are B; therefore Q B are A" is valid with Q = "some" or Q = "all". Even if she doesn't have a special-purpose inference rule for dealing with this sort of situation, Mrs C might still tackle the problem in a logical way, by asking herself if the meanings of "some" and "all" license transposition of their arguments. If this is how she approaches the task, it is clear that mistakes might arise, and given that most quantifiers are symmetric (and therefore license conversion) and that, presumably, people have a general preference for symmetry, it is to be expected that Mrs C should be more prone to make mistakes with "all" than with "some".[14]

This example shows two things. First, it illustrates that the Tarskian view on logical reasoning offers a new perspective on reasoning errors. Secondly, it shows that some errors are plausibly viewed as genuinely logical. On the account just sketched, illicit conversion is precisely the kind of error people would commit if they reasoned logically, and it is hard to see how a non-logical account could accommodate this line of explanation, simply because it hinges on the assumption that, occasionally, people misidentify the logical form of an argument. If this much is correct, illicit conversion is a logical error not only in the sense that it violates the logic of certain quantifiers but also in the sense that it arises because people reason logically.

This is not an isolated example. Logical errors in the sense just explained also occur in other types of syllogistic reasoning, for instance. We saw before that the most widely endorsed syllogisms tend to be valid. But logical validity also has an effect on errors in syllogistic reasoning, in that high-frequency errors tend to occur with invalid arguments that resemble valid ones. For example, of the four syllogisms which have three occurrences of the quantifier "all", only one is valid: "all B are C, all A are B; therefore all A are C". In the vast majority of cases, this argument is recognised as valid; in fact, it is one of the easiest syllogisms. But the other three straight-"all" syllogisms appear to share in this popularity, and are endorsed well above average. For instance, according to Chater and Oaksford's (1999) data, which combine the results of five experiments conducted by other researchers, these three syllogisms were endorsed

[14] See Geurts (2002, 2003) for more discussion of symmetry in interpretation and inference.

63% of the time on average ($P_{.9}$ = 56.5). As in the case of illicit conversion, this evidence suggests rather strongly that people may get the logical form of an argument wrong because it resembles the form of a valid argument. If so, they reason logically, and their errors show it.

One of the recurrent themes in criticisms of mental logic is that, in general, logic-based approaches have difficulties explaining reasoning errors. In this section I have tried to show, first, that this criticism is not justified, and secondly, that certain errors documented in the literature actually indicate that reasoning is at least in part a logical process.

5 Logical content

So-called "content effects" have played an important part in the reasoning literature, and it is widely believed that they pose a serious threat to logicist theories of inference:

> One problem is that the content of premises can have striking effects on the conclusions that subjects draw from them [...]. Yet formal rules by definition are supposed to apply regardless of content. (Johnson-Laird, Byrne, and Tabossi 1989: 659)

Claims to much the same effect will be found in practically any critical review of logic-based theories. There are a number of problems with this line of criticism, beginning with the very notion of content effect. Johnson-Laird et al. discern an opposition between reasoning on the basis of content, on the one hand, and drawing inferences by means of "formal rules", on the other; and most criticisms of mental logic rely on the same contrast. However, despite its popularity, this opposition is a false one. The validity of "p and q, therefore p", for instance, is determined almost entirely by the semantic content of "and", and if a "formal rule" is used to capture this inference, it can hardly be said to apply regardless of content.[15] More generally, semantic content is one

[15] There is of course a sense in which mental rules of inference are irredeemably formal: they perform operations on representations without having direct access to what these representations stand for. But in this sense, *all* psychological theories of reasoning are formal, so I take it that this notion of formality is not at issue.

of the crucial ingredients in the concept of logical validity, as we have seen, so the widespread belief that logical inference is content-free is a contradiction in terms.

Instead of saying that logical reasoning is blind to the content of a statement, it is better to say that it is contingent on certain aspects of its content only: logical reasoning is oblivious to non-logical content. So, properly speaking, in order to demonstrate content effects it would have to be established that reasoning processes fail to differentiate between logical and non-logical content. And this is not a trivial task, if only because that distinction anything but cut-and-dried. To illustrate, let us briefly dip into what is certainly the richest source of content effects thus far, Wason's (1966) selection task.[16] In the original version of the Wason task, subjects are presented with four cards drawn from a pack which all have a letter on one side and a number on the other, and the subjects' task is to verify a conditional statement like:

(1) If a card has an A on one side, it has a 7 on the other side.

Two of the facing sides of the four cards each match one half of the conditional rule, whereas the other two don't, like this: "A K 4 7", and subjects are to say which cards have to be turned over in order to determine whether or not (1) holds. Typical choices in this situation are either only the A-card or that and the 7-card; the correct response (A and 4) is rarely observed.

From the early seventies onwards, it has become increasingly clear that it matters a great deal what the selection task is about. In particular, it has often been found that if the topic is a deontic one, the task is greatly facilitated. For example, when Griggs and Cox (1982) used the rule:

(2) If a person is drinking beer, then that person must be over 19 years of age.

subjects' performance was much better than on the original task. Such findings are widely regarded as content effects, the (usually tacit) assumption being that logical form remains the same across experiments. That is to say, it is assumed that the conditional statement subjects have to evaluate always is of the form "if p then q". In deontic tasks, the consequent of the conditional happens to contain a modal statement, but this is

[16] See Stenning and van Lambalgen (2001) for a recent overview and extensive discussion.

regarded as a difference in subject matter rather than logical form. This assumption has recently been challenged by Fodor (2000).[17]

Fodor's claim is not just that deontic versions of the selection task use conditionals of the form "if p then must q", while the standard version does not. On Fodor's account, the difference goes deeper, because the logical form of "if ... then" statements like (2) is not that of a conditional in the first place. Whereas (1) asserts that there is a conditional relationship between a card's having an A one side and its having a 7 on the other side, (2) unconditionally asserts that people who drink beer must be over 19 years of age. In the latter case, it is the function of the if-clause to restrict the range of individuals that fall under the edict. If this is so, then it is unremarkable, according to Fodor, that subjects find it easy to see that an 17-year old is a potential violator of (2), for the sentence is used to state that everybody who drinks beer must be over 19 years of age.

Although it remains to be seen whether Fodor's analysis of conditionals is correct, it is of interest here chiefly for the strategy it pursues. Whereas the difference between (1) and (2) is usually explained in terms of non-logical content, Fodor's suggestion is the variation in the data is better explained "... by recognizing a difference in the *logical form* of declarative and deontic hypotheticals." (Fodor 2002: 221; emphasis added). Whether or not it is correct, this manoeuvre is a legitimate one, because differences between deontic and non-deontic statements have been observed across many different versions of the selection task, and it is not necessary for two constructions of the same grammatical structure to have to same logical form. But by the same token, it illustrates how difficult it can be to draw a boundary line between content effects and variation in logical form.

One influential attempt at explaining the content effects found in the Wason task is Cheng and Holyoak's theory of "pragmatic reasoning schemas".[18] According to Cheng and Holyoak (1985: 394) "... the view that people typically reason in accord with formal logic has been overwhelmingly refuted by evidence based on experiments in conditional reasoning," and therefore a non-logical approach is called for. The alternative view advocated by Cheng and Holyoak is that human reasoning relies on

[17] For further discussion, see Beaman (2002), Fodor (2002).

[18] Cheng and Holyoak (1985, 1989), Cheng, Holyoak, Nisbett, and Oliver (1986).

pragmatic reasoning schemas, such as the permission schema, which consists of the following rules:

> Rule 1: If the action is to be taken, then the precondition must be satisfied.
> Rule 2: If the action is not to be taken, then the precondition need not be satisfied.
> Rule 3: If the precondition is satisfied, then the action may be taken.
> Rule 4: If the precondition is not satisfied, then the action must not be taken.

As applied to the conditional statement in (2), Rule 1 implies that persons drinking beer should be checked, and Rule 4 urges us to check underage drinkers. But as the subject matter of the statement in (1) doesn't fall under the permission schema, the schema doesn't apply in this case. Thus it is correctly predicted how subjects respond to deontic statements such as (2), and that (1) is different from (2).[19]

Cheng and Holyoak adduce a number of considerations to show that their approach is "qualitatively different" (1985: 395) from other theories of deductive reasoning, especially logical ones. None of their arguments withstand closer scrutiny, however. For example, they observe that their pragmatic reasoning schemas don't contain the schematic letters that are characteristic of inference rules: "The permission schema ... contains no context-free symbols such as *p* and *q* above." (1985: 396) However, if variables are absent from Cheng and Holyoak's rules it is just because their formulation remains at an informal level. If one had to spell out in more explicit terms what (say) Rule 1 is meant to express, p's and q's could not be dispensed with (and they duly appear in Evans et al.'s 1993 discussion of permission schemas). Spelling matters aside, Cheng and Holyoak insist that their permission schema is *abstract* in the sense that it holds for any type of action, which is to admit in so many words that variables are required. Another consideration adduced by Cheng and Holyoak is that their permission schema is not equivalent to the material conditional in standard propositional logic (1985: 397). This much is correct, of course, but it doesn't even begin to show that permission schemas are non-logical. In short, there is no reason why Cheng and Holyoak's permission schema couldn't be construed as a *logic* of permission statements.

---

[19] Hence, permission schemas do not a provide a complete account of performance on the Wason task, but only of the main content effect that has been found in this task.

Cheng and Holyoak's permission schema is supposed to capture (part of) the intuitive meanings of deontic words such as "must" and "may", and when it is applied everything but these meanings is abstracted over. So the long and short of it is that there is no principled difference between Cheng and Holyoak's permission schema and standard logical treatments of modal operators, or any other logical constants, for that matter. Moral: in so far as that Cheng and Holyoak's proposal is on the right track, it confirms that people reason logically. Claims to the contrary are based on misunderstandings concerning the notion of content and, consequently, on an overly narrow conception of what logical reasoning is.

Let us suppose, contrary to fact, that the distinction between logical and non-logical content is always clear-cut, and that, moreover, it has been established that performance on a given task is influenced by variations in the non-logical parts of the materials used. Would that show that the kind of reasoning employed in this task is not logical? It would not. What it would show is that the reasoning processes subjects engage in are not *merely* logical. This is a rather obvious point, I admit, and if I bring it up nonetheless it is only because it is not always duly appreciated (as witness, for example, the remark by Johnson-Laird et al. quoted at the beginning of this section). If someone were to claim that human reasoning is a deterministic affair based entirely on logical principles, then, and only then, any effects of non-logical content would count as negative evidence. But surely such a claim is too radical even for the staunchest logicist. As I noted in the introduction, most people working on the psychology of reasoning would say that reasoning is a hybrid process, and if this much is accepted the mere existence of content effects will not decide the fate of any theory.

Thus far my line of argument has been defensive: I have been trying to expose various ways that content effects cannot be used to argue against logicist theories of reasoning. But there is also a more constructive way of framing the content issue. If the Tarskian view on reasoning is on target, people should at least sometimes be sensitive to the logical form of an argument. *To the extent that* reasoning is a logical process, it is indifferent to non-logical content. What we should find, therefore, is data showing that by and large or in certain respects performance is not affected by variations in non-logical content. Is there any experimental evidence that fits this bill? Yes, there is, and plenty of it, too. One example is the Wason task, which has been varied in a huge number of ways, showing time and again that, in order to verify a rule of the form "if p then q", subjects nearly always select the p-card. This is in line with evidence from

many studies showing at people are extremely good at modus ponens, regardless whether the subject matter is letters, people, fish, cards, machines, or whatever.[20] And it is the same with other propositional arguments and a good many arguments involving quantified statements, especially syllogisms. Granted, the experimental evidence for (relative) content-independence is less abundant in these cases than it is for conditionals, but presumably that is because it seems so unlikely that negative evidence could be found. It would be nice to have experimental results demonstrating that people endorse "p and q; therefore p", no matter what p and q stand for; but it is fairly obvious why nobody should have bothered.

6 Whence mental logic?

The upshot of the foregoing discussion is that to some degree at least reasoning is a logical process. But where did our mental logic come from? Is it something an individual acquires during its lifetime, or has our species (like other species, perhaps) acquired it in the course of evolution? Either way, it has been claimed, logicist theories of reasoning are in trouble; for it is obscure how something like a mental logic could ever develop, be it on the level of the phenotype or that of the genotype.

> The acquisition of deductive competence is profoundly puzzling for theories based on formal rules. Given some existing deductive ability children might learn other rules of inference by generalization ..., but how could they learn formal rules in the first place? ... So intractable is the problem for formal rules that many theorists suppose that deductive ability is not learned at all. It is innate. Fodor (1980) has even argued that, in principle, logic could not be learned. The difficulty with this argument is not that it is wrong, although it may be, but that it is too strong. It is hard to construct a case against the learning of logic that is not also a case against its evolution. (Johnson-Laird and Byrne 1991: 204)[21]

[20] See e.g. Evans (1977), Marcus and Rips (1979), Braine, Reiser, and Rumain (1984), Markovits (1988), Johnson-Laird, Byrne, and Schaeken (1992), Thompson (2000).

[21] For more discussion in the same critical vein, see Johnson-Laird (1983), Cosmides (1989), Cosmides and Tooby (1992), Bara, Bucciarelli, and Lombardo (2001).

This is partly off the mark: the scope of Fodor's argument is much broader than Johnson-Laird and Byrne allow for (as indeed Fodor's title indicates), and if the argument is correct it equally affects all theories of reasoning, Johnson-Laird and Byrne's included.[22] But even so their worry seems justified, though it is not easy to pin down the problem exactly. The heart of the matter, I believe, is the following. On the common proof-theoretical conception, mental logic is a collection of inference rules akin to a system of natural deduction. Mental logic is conceived of as an autonomous faculty in the mind, and as a general-purpose tool, ready to be deployed for any task at all. It is certainly not a bad thing for a mental module to be autonomous and domain-independent, but these virtues become somewhat of an embarrassment when we are to explain how mental logic could be acquired. And passing the buck to Mother Nature will not do, as Johnson-Laird and Byrne rightly point out, because if we can't explain how mental logic is acquired, the prospect of explaining its evolution is no brighter.

How then are we to respond to Johnson-Laird and Byrne's charges? In part, by insisting that the theory of mental logic is entitled to a measure of nativism. It is more than likely that certain elementary notions of logic are part of our genetic endowment. For example, it seems quite plausible to me that some of the concepts formalised in the propositional calculus are innate: negation, conjunction, disjunction, conditionality. Not of course in the sense that we come equipped with a full-blown truth-functional notion of negation (say), but in the sense that the truth function has an intuitive precursor which is so basic that it could hardly have been acquired by learning.

A modicum of nativism lessens the impact of Johnson-Laird and Byrne's objection, but it is not the whole story. Consider the following piece of reasoning, for example: "Jones got more than 75% of the votes; therefore, Jones got more than 50% of the votes." This argument is logically valid: not to put too fine a point on it, knowing the meaning of "more than n %" will suffice to appreciate that the premiss necessitates the conclusion. But although it is plausible enough that the concept expressed by "more than" is innate, the same cannot be said of natural numbers (save perhaps for the lower two or three), let alone percentages. It is apparent, then, that logical reasoning is not just a matter of applying innate knowledge.

Once again, I think the problem arises from a too narrow view on logic. In this case, the presupposition that, by definition, a logic is a system of rules, and general-

[22] See Laurence and Margolis (2002) for in-depth discussion of Fodor's puzzle, as they call it.

purpose rules at that, fosters the notion of mental logic as an autonomous faculty of the mind. But in a Tarskian perspective there is no need to conceive of mental logic as a separate module, much less as an autonomous one. Someone may be said to reason logically if he recognises that a conclusion follows necessarily from a set of premisses simply by virtue of (what he takes to be) their logical form, where, crucially, logical form is not something that is inherent in an argument, but merely something that emerges when parts of it are abstracted out (see Sections 2 and 3 above). The key ingredients of logical reasoning, thus understood, are necessity and a form of abstraction, both of which may be assumed to be available on independent grounds already. Logical reasoning is unlike any other mode of inference people engage in, but it doesn't require abilities that are specifically logical. This, I take it, is as we should like it to be, and it follows from the Tarskian view on mental logic that I am advertising.

To return to the topic of the present section, I have urged that we can reasonably speak of mental logic as a unique way of dealing with inferential tasks, without implying the existence of an autonomous faculty of the mind. Logical reasoning is like writing poetry: as an activity it is unlike any other, and some people are more adept or better at it than others, but the kind of abilities it requires are not unique to it. Hence, logical reasoning does not present a profound problem for theories of acquisition and/or evolution unless poetical skill does too.

If the foregoing argument remains at fairly abstract level, it can be fleshed out by reverting to a theme introduced earlier, of the connection between interpretation and inference.[23] I have suggested that people may tackle some reasoning tasks by concentrating their attention on one or two of the expressions it contains and their meanings. For example, in order to go from "Barney is in Florence and Fred is in Berlin" to "Barney is in Florence" we only need to note that the premiss contains a token of "and" and that its first conjunct is identical with the conclusion. Together with the meaning of "and", this is enough to realise that the inference is valid. Rules are not required, nor need we worry about the meanings of the conjoined sentences or their parts. This is, I believe, an entirely plausible story about how people actually deal with such inferences, but whether or not this particular analysis is right, what is important is that the Tarskian view on mental logic licenses this type of explanation, and therewith

[23] See sections 3 and 4 above. The importance of semantics to the psychology of reasoning has previously been stressed by Stenning and van Lambalgen (2001) and Geurts (2003).

suggests how logical reasoning may be acquired. For, in so far as logical reasoning can be accounted for along these lines, we merely need to explain how people manage to represent the meanings of certain expressions, and how they acquire the ability to focus on part of the available information to the exclusion of all the rest. Thus the problem of the acquisition of mental logic reduces to two issues that, though by no means trivial, must be dealt with anyhow — a far cry from the "profound puzzle" of Johnson-Laird and Byrne's statement.

References

Bara, B.G., M. Bucciarelli, V. Lombardo 2001: Model theory of deduction: a unified computational approach. *Cognitive Science* 25: 839-901.

Beaman, C.P. 2002: Why are we good at detecting cheaters? A reply to Fodor. *Cognition* 83: 215-220.

Braine, M.D.S. and B. Rumain 1983: Logical reasoning. In: J.H. Flavell and E.M. Markman (eds.), *Handbook of Child Psychology, Volume 3: Cognitive Development*. Wiley and Sons, New York. Pp. 263-340.

Braine, M.D.S., B.J. Reiser, and B. Rumain 1984: Some empirical justification for a theory of natural propositional logic. In: G.H. Bower (ed.), *The Psychology of Learning and Motivation*. Academic Press, New York.

Braine, M.D.S. and D.P. O'Brien 1991: A theory of *if*: a lexical entry, reasoning program, and pragmatic principles. *Psychological Review* 98: 182-203.

Chater, N. and M. Oaksford 1999: The probability heuristics model of syllogistic reasoning. *Cognitive Pyschology* 38: 191-258.

Cheng, P.W. and K.J. Holyoak 1985: Pragmatic reasoning schemas. *Cognitive Psychology* 17: 391-416.

Cheng, P.W. and K.J. Holyoak 1989: On the natural selection of reasoning theories. *Cognition* 33: 285-313.

Cheng, P.W., K.J. Holyoak, R.E. Nisbett, and L.M. Oliver 1986: Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology* 18: 293-328.

Cosmides, L. 1989: The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition* 31: 187-276.

Cosmides, L. and Tooby, J. 1992: Cognitive adaptations for social exchange. In: J.H. Barkow, L. Cosmides, and J. Tooby (eds.), *The Adapted Mind: Evolutionary psychology and the Generation of Culture*. Oxford University Press. Pp. 163-228.

Dickstein, L.S. 1981: Conversion and possibility in syllogistic reasoning. *Bulletin of the Psychonomic Society* 18: 229-232.

Etchemendy, J. 1990: *The Concept of Logical Consequence*. Harvard University Press, Cambridge, Mass.

Evans, J.S.B.T. 1977: Linguistic factors in reasoning. *Quarterly Journal of Experimental Psychology* 29: 297-306.

Evans, J.S.B.T., S.E. Newstead, and R.M.J. Byrne 1993: *Human Reasoning: The Psychology of Deduction*. Lawrence Erlbaum, Hove.

Fodor, J. 1980: Fixation of belief and concept acquisition. In: M. Piattelli-Palmarini (ed.), *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*. Harvard University Press, Cambridge, Mass. Pp. 143-149.

Fodor, J. 2000: Why we are so good at catching cheaters. *Cognition* 75: 29-32.

Fodor, J. 2002: Reply to Beaman. *Cognition* 83: 221.

Ford, M. 1995: Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition* 54: 1-71.

Geurts, B. 2002: Donkey business. *Linguistics and Philosophy* 25: 129-156.

Geurts, B. 2003: Reasoning with quantifiers. Forthcoming in *Cognition*.

Griggs, R. A. and J.R. Cox 1982: The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology* 73: 407–420.

Hanson, W.H. 1997: The concept of logical consequence. *The Philosophical Review* 106: 365-409.

Johnson-Laird, P.N. 1975: Models of deduction. In: R.J. Falmagne (ed.), *Reasoning: Representation and Process in Children and Adults*. Erlbaum, Hillsdale, NJ.

Johnson-Laird, P.N. 1983: *Mental Models*. Cambridge University Press.

Johnson-Laird, P. N. 1997: Rules and Illusions: A critical study of Rips's The Psychology of Proof. *Minds and Machines* 7: 387–407.

Johnson-Laird, P. N., R.M.J. Byrne, and P. Tabossi 1989: Reasoning by model: The case of multiple quantification. *Psychological Review* 96: 658–673.

Johnson-Laird, P. N. and R.M.J. Byrne 1991: *Deduction*. Lawrence Erlbaum, Hillsdale, NJ.

Johnson-Laird, P.N., R.M.J. Byrne, and W. Schaeken 1992: Propositional reasoning by model. *Psychological Review* 99: 418-439.

Laurence, S. and E. Margolis 2002: Radical concept nativism. *Cognition* 86: 25-55.

Lehman, D. R., R.O. Lempert, and R.E. Nisbett 1988: The effects of graduate training on reasoning: Formal discipline and thinking about everyday-life events. *American Psychologist*, 431-442.

Marcus, S.L. and L.J. Rips 1979: Conditional reasoning. *Journal of Verbal Learning and Verbal Behavior* 18: 199-224.

Markovits, H. 1988: Conditional reasoning, representation, and empirical evidence on a concrete task. *Quarterly Journal of Experimental Psychology* 40A: 483-495.

Newstead, S.E. 1989: Interpretational errors in syllogistic reasoning. *Journal of Memory and Language* 28: 78-91.

Nisbett, R.E., G.T. Fong, D.R. Lehman, and P.W. Cheng 1987: Teaching reasoning. *Science* 238: 625-631.

Oaksford, M. and N. Chater 2001: The probabilistic approach to human reasoning. *Trends in Cognitive Sciences* 5: 349-357.

Osherson, D.N. 1977: Natural connectives: a Chomskyan approach. *Journal of Mathematical Psychology* 16: 1-29.

Ray, G. 1996: Logical Consequence: A defense of Tarski. *Journal of Philosophical Logic* 25: 617-677.

Rips, L.J. 1983: Cognitive processes in propositional reasoning. *Psychological Review* 90: 38-71.

Rips, L.J. 1994: *The Psychology of Proof: Deductive Reasoning in Human Thinking*. MIT Press, Cambridge, Mass.

Sher, G.Y. 1996: Did Tarski commit "Tarski's fallacy"? *The Journal of Symbolic Logic* 61: 653-686.

Smith, E.E., C. Langston, R. Nisbett 1992: The case for rules in reasoning. *Cognitive Science* 16: 1-40.

Stenning, K.E. and M. van Lambalgen 2001: Semantics as a foundation for psychology: a case study of Wason's selection task. *Journal of Logic, Language and Information*10: 273-317.

Tarski, A. 1936 (1956): On the concept of logical consequence. In: *Logic, Semantics, Metamathematics*. Oxford University Press. Pp. 409-420. Translation of a paper first appeared in Polish in 1936.

Thompson, V.A. 2000: The task-specific nature of domain-general reasoning. *Cognition* 76: 209-268.

Wason, P. C. 1966: Reasoning. In: B. Foss (ed.), *New Horizons in Psychology*. Penguin, London. Pp. 135–151.

Wilkins, M.C. 1928: The effect of changed material on ability to do formal syllogistic reasoning. *Archives of Psychology* 16: 1-83.