2

Why It Took So Long to Bake the Mental-Logic Cake: Historical Analysis of the Recipe and Its Ingredients

Luca Bonatti
Laboratoire des Science Cognitives et Psycholinguistique

> *. . . at least to me, the use of technically advanced machinery in analyzing reasoning is encouraging; after all, Aristotle thought about reasoning; one would like to see clearly what one has that he did not have! (It is no comfort to know that over 2000 years have passed since his time unless one sees just how one has used the experience of these 2000 years.)*
>
> —Kreisel (1967, p. 271)

"Enough with mental logic! It has been around for three millennia, and no theory came out of it. It's a dead research program!" You certainly must have happened to overhear speech streams very much like this. (They generally continue with the utterer praising the virtues of sexier alternatives, such as models, interconnected networks, space phases, or what have you.) In case you wanted to know what to respond, I can help you. The right response is, "Yes and no."

However, "yes and no" is a contradiction and you will not be happy with it, because you, like me and everybody else, possess a natural logic and natural logic abhors contradictions. The *real* right answer is, "In one sense yes, and in another sense, no." The "In one sense, yes" part is easy: It is a fact that the idea that we have a logic in our mind goes back 3,000 years or so. The "in another sense, no" part is a bit more complicated. I tell a long story about how certain ideas have to be blended with the right ingredients, how the ingredients have to be carefully mixed together, and how the mixture must be stored in a warm place far from drafts in order for the yeast to grow. This paper is a memo to recall the basic steps for the recipe in order to be successful.

BLEND LOGIC WITH PSYCHOLOGY, BUT NOT TOO MUCH: WHEN LOGIC AND PSYCHOLOGY WERE ONE

Yes, something like the idea that there is a logic in our mind has been floating around for millennia. But could it actually develop into a real psychological thesis? Have the cake ingredients been available all along? Consider a step as trivial as this: If you want to claim that there is a relation between logic and our psychology, first, you had better be clear about the difference between logic and psychology, and, second, you had better get it right.

Yet this has not been a step easy to achieve, because the mental logic thesis was almost always meant in a very strong sense—as the thesis that logic *is* the science of the laws of thought. As a consequence, logic was considered *part of* psychology. This inclusive thesis remained surprisingly stable across the centuries, although it was clearly stunting its development by cutting its ties with the rest of mathematics.[1]

To a certain extent, however, this stability should not be a surprise. As Kant noticed (1781/1966, pp. 17–18), logic also remained substantially unchanged for centuries, as if all possible forms of reasoning had been laid down once and for all by Aristotle and his medieval commentators. So, for a long time there was no reason to change the doctrine of the psychological nature of logic, because logic was stable because considered complete and psychology was stable because non-existent.

What begins to be surprising is that this thesis resisted even when the horizons of logic broadened, for the first time with the work of Boole. Boole first distinguished propositional and predicative logics, and first noticed a central element for all future developments, namely, the formal nature of logic.[2] Yet, even if he dared to unrest its millennial stability,

[1]The fact that logic developed so late if compared with the rest of mathematics is indeed a puzzle. The puzzle is even deeper if one think that little or no mathematics was necessary to develop it, and that the necessary mathematics existed centuries before logic started its own solid development. I am suggesting that part of the explanation has to be found in the fact that the pie was cut in the wrong way. Logic was just not considered a discipline of the same nature as mathematics. The strong inclusive version of the mental-logic hypothesis bears part of responsibility for the anomalous development of logic.

[2]So also in Mangione (1993): "What does the revolutionary nature of Boolean logic consist in? . . . Boole clearly and definitely states the *formal* nature of the calculus in general, in the sense that also the formulation of a logical calculus is a formal construction to which interpretation is added from outside and does not constitute any longer the exclusive and primary basis from which the formal structure is abstracted" (pp. 106–107).

Boole did not abandon the idea that an investigation of logic is at the same time an investigation of "the fundamental laws of those operation of the mind by which reasoning is performed" (1854, p. 1). Moreover, Boole also made free use of another traditional ingredient that gives scarce results when added to the mental-logic cake. He considered logic as a discipline *sui generis* whose laws are not to be discovered by empirical inspection, like all other sciences require, but can be known entirely by introspection. This conception may be perfectly adequate for logic—after all, this is what intuitionism claims—but obviously cannot work for psychology. Yet, because for Boole logical laws are laws of thinking, a consequence of the mixture of the two traditional ingredients was that for him the "science of the mind"—what we would today call the *psychology of reasoning*—need not be grounded in experience:

> The general laws of nature are not, for the most part, immediate objects of perception. . . . On the other hand, the knowledge of the laws of the mind does not require as its basis any extensive collection of observations. The general truth is seen in the particular instance, and it is not confirmed by the repetition of instances. (1854, p. 4)

Let me stress that for centuries this view, according to which in order to achieve "knowledge of the mind" you need only to introspect your own mind, was never seriously questioned. Yet, it is as plausible as the view that you can discover the laws of your natural language parser by introspection; psycholinguistics would not go that far with the latter, nor the psychology of reasoning with the former.

Even if his view of mental logic was so much indebted to the past, Boole's (1854) analysis did break with tradition in an important point: He introduced the idea that the "laws of thought" have a mathematical character, more precisely algebraical. And this agreement between algebra and logic of thought, he conceded, must be found a posteriori:

> There is not only a close analogy between the operations of the mind in general reasoning and its operations in the particular science of algebra, but there is to a considerable extent an exact agreement. . . . Of course the laws must in both cases be determined independently; any formal agreement between them can only be established a posteriori by actual comparison. (p. 6)

Boole also speculated that because the *content* of number theory and logic are different, the "agreement" between operations of the mind and algebra must concern only the *processes* involved. Mental processes and algebraic processes have something in common. What, though, and why?

He did not give a full answer. Perhaps an answer cannot be given, he suggested, because it lays beyond our intellectual capabilities:

> Whence it is that the ultimate laws of logic are mathematical in their form, . . . are questions upon which it might not be very remote from presumption to endeavor to pronounce a positive judgment. Probably they lie beyond the reach of our limited faculties. (1854, p. 11)

This profession of necessary ignorance was too pessimistic. There was a possible explanation of what unifies mathematics and thinking. Boole himself had remarked insightfully that reasoning is a process of compositions of symbols according to laws, that symbols are arbitrary signs, and that their rules of composition are common to both logic and the science of numbers. In order to see the full answer, however, he needed some other ingredients. The missing parts of the recipe were a foundational theory of mathematics that would stress the centrality of signs and combinations of signs, and a psychologically plausible theory that would assign to signs a key role in our thought processes. Both were to be provided in our century, but only, paradoxically, when philosophers and logicians had already abandoned the program of mental logic.

BUT NOT TOO LITTLE EITHER: LOGICIANS ABANDON PSYCHOLOGY . . .[3]

The first dissonant tune was sung by Frege and Russell. With their works, for the first time logic became theoretically separated from psychology. This is indeed good news, if I am right in thinking that confusion about the boundaries between the two disciplines was slowing the development of both. The trouble is, they got separated too much. Logic divorced psychology only to find itself married with metaphysics, and, in order to be accepted in the new family logicians, gave up any ties with the investigation of thinking processes.

Frege (1979) argued forcefully not only that logic—the science of thought, in the special sense he was conceiving it—is separated from psychology—for him, the science of thinking—but also that the latter is the mortal enemy of the former. In his colorful terms:

> [I]t is the business of the logician to conduct an unceasing struggle against psychology and those parts of language and grammar which fail to give untrammeled expression to what is logical. (p. 6)

[3]Macnamara (1986, pp. 12–20), to which the reader is also referred, makes much the same point I am arguing in this section.

> Since thoughts are not mental in nature, it follows that every psychological treatment of logic can only do harm. It is rather the task of this science to purify logic of all that is alien and hence of all that is psychological. (pp. 148–149)

The reasons motivating his calls for purification are interesting for us because they show that Frege was moved by his deep anti-idealism[4] rather than by an opposition to the mental logic hypothesis *qua* psychological hypothesis. If we fail to separate logical and psychological content, Frege argued, we obliterate the difference between the strictly subjective and noncommunicable ideas and the universal logical principles, so nothing would guarantee that two persons can entertain *the same* content, asserting it or denying it. As a consequence, a contradiction between two persons would be impossible, and with it communication as well as science. There would be no way to explain the notion of truth, idealism could not be refuted, solipsism would be inescapable, everyone would forever be condemned to live in his or her own private world.[5]

Even for Russell (1919)—and even more directly—logic became an instrument for the discovery of the correct ontology, a position pushed to its extreme consequences by his claim that logic should not include unicorns any more than zoology does (see 1919, pp. 169–170). As Di Francesco (1991) commented, "for Russell 'Logic' does not mean anything linguistic or formal in nature, in our sense. Quite the contrary, what he was looking for was not a set of syntactic rules, but rather the correct description of the objective relations between entia or propositions, considered as complex groups of entia" (p. 87).

Russell and Frege were not alone. For different reasons the logical neopositivists also accepted the new dogma. They saw in the novel developments of logic an occasion to explain at the same time the necessity of mathematical truth—a real mystery for traditional empiricism—and the construction of complex experiences: Mathematics is necessary because it is as empty as logic, and experience is structured in public objects because objects are a logical composition of sense data. Notice that in both roles logic takes over functions that have no connection whatsoever with psychology or the mind. This was somewhat a forced choice for the neopositivists. They were trying to play cards on two separate tables. They wanted both to hold that the origin of meaning is entirely grounded in experience, and to free experience from any connotation of privacy. As

[4]"Idealism" here does not include transcendental idealism, but the mixture of psychologism and post-Hegelianism that was current in Germany and England at the end of the century.

[5]See Frege (1979); for example, in "Logic," and "The Thought." See also Macnamara (1986), pp. 14–16.

a result, to their ears any talk of inner rules of thought or of mental entities appeared suspect, and behaviorist or reductionist conceptions of the mind sounded more appealing. In America, where the neopositivists migrated to escape Nazi fascism, the climate was perfect for these ideas to flourish. So even Quine, the philosopher who killed the neopositivist program with his criticisms of the analytic–synthetic distinction, still maintained their radical skepticism about the mind and their wary attitude towards any use of logic for psychological purposes.

In short, among the large majority of philosophically minded logicians, showing interest in psychological processes became a sort of behavior that well-mannered people should avoid. Logic and reasoning took different routes precisely when logic was undergoing an impressive development, and precisely under the influence of the pivotal figures of this development.

However, this turn was motivated less by any substantial arguments, or by any problems with the psychological feasibility of mental logic, than by a change in the general cultural climate. This is clear especially in Frege and Russell, whose statements about the separation between logic and psychology can be seen as an epiphenomenon of the general rebellion against German and English idealism from which 20th-century analytic philosophy stemmed,[6] but it is worth recalling that in the first decades of the century many other philosophers of different orientations, such as Moore, Meinong, and Husserl, attacked idealism in one way or another.[7]

So the antipsychological shift really was radical. However, it is important to realize that even conceding to Frege and the anti-idealists that the logic of thought is not the logic of thinking, there still is the possibility of having *both*: a metaphysical theory of thought *and* a theory of logical thinking, in which some of the logical rules discovered in the investigation of thought have a causal role in our thinking processes. For some reason, Frege and the post-Fregeans never conceived of this intermediate possibility.

So, although of extreme historical importance, the rejection of the mental-logic hypothesis by the new logicians was not really motivated. In the past, logicians had made the mistake of including logic within psychology and had made the further mistake of assuming the mental-logic point of view without giving any substantial argument for it. The new logicians corrected the first mistake by sharply separating logic from psychology. However, in one sense they made the same basic second mistake already made by their predecessors: They took a stand towards

[6]See Dummett (1981) for Frege, and Di Francesco (1991) for Russell.

[7]Macnamara (1986) makes this point especially with respect to Husserl's view of logic, which drastically changed under the influence of Frege. But Husserl was no special case: A plethora of authors was coming to the same conclusions in the same years.

the mental-logic program—a negative one in this case—without advancing any real argument.

. . . AND PSYCHOLOGISTS ABANDON LOGIC

Nevertheless, for independent reasons the same ideas were gaining the favor of experimental psychologists and became the standard view at least until the early 1960s. Many factors would probably need to be analyzed to explain the success of this attitude, but a major one is to be found in an unusual alliance of interests between otherwise opposite approaches. Notwithstanding the gulf separating behaviorism and cognitive psychology, at least for a certain period both schools held the same critical position about mental logic. Behaviorists were led to it by their general skepticism about talk of inner mental processes, including inner mechanisms of reasoning. In the limited interest shown for the problem, behaviorism seemed to have its own way to explain the "logical" appearance of people's responses. Problem solution was considered a learned drive, a kind of associative behavior. To the objection that thinking did not really look like a type of behavior, the answer was that it is in effect an evolutionary sophistication of action, a sort of delayed action that does leave silent behavioral traces in the form of implicit responses whose behavioral realization is only more difficult to be detected, such as subvocal talk or other muscular movements that repeat the full normal behavioral response at a smaller scale.

The underlying idea was that a "central controller" (of which mental logic is a partial description) is not needed for mental tasks: Ultimately, those can be traced back to peripheral muscular activities. So Watson (1924) wrote:

> The behaviorist advances the view that *what the psychologists have hitherto called thought is in short nothing but talking to ourselves.* . . . My theory does hold that the muscular habits learned in overt speech are responsible for implicit or internal speech (thought). (pp. 14–15, italics added)[8]

Skinner later rejected the subvocal theory of thought, but not the idea that thinking and reasoning can be reduced to behavior. If it was impossible to explain it away by means of a single correlated behavior, for him thinking could be seen as an emergent property of complex behaviors. The basic point remained the same: Thinking is not a succes-

[8]See also Humphrey (1963), pp. 185–216.

sion of internal states, but a far consequence of environmental variables on the organism.[9]

The need to dispose of mental vocabulary—and with it, by instantiation, mental logic—looks like a necessary corner into which the behaviorists had to paint themselves. Well, one can say, too bad for the behaviorists. However, paradoxically, even the newborn cognitive psychology, which had no need to put itself in such an uncomfortable position, assumed roughly the same attitude toward mental logic. Although the "New Look" psychology rehabilitated a vocabulary of inner mental processes and flows of information, initially it focused research on the plasticity of mental processes, on the interaction and the mutual influence of reasoning and perception, and on the strong effect of context on both. To this general picture of the functioning of mind the "rigidity" of logical rules was unappealing, whereas a more pragmatic view of reasoning was better suited. Thus we find what is considered the first bible of cognitive science stating that "much of human reasoning is supported by a kind of thematic process rather than by an abstract logic. The principal feature of this thematic process is its pragmatic rather than its logical structure" (Bruner, Goodnow, & Austin, 1956, p. 104). Once again, the intuition was correct but not its implicit conclusion. That pragmatics influences reasoning (as well as other mental processes) is obviously true. However, it is one thing is to say that a theory of reasoning has to take into account pragmatic factors, and another to say that *because* reasoning is influenced by pragmatic factors, it has no *formal* components and no mental logic rules. Compare the following: It is one thing is to say that language interactions among humans are heavily influenced by semantics, and another to say that *because* semantics has a great influence in language, there is no syntax. In both cases the former thesis does not imply the latter. New Look psychologists missed this difference and thus, even if they had no a priori reasons to reject mental logic, they treated it almost as badly as their behaviorist adversaries did.

ADD A GOOD MEASURE OF FORMALISM: HILBERT AND THINKING AS A SYMBOLIC PROCESS

Ideologically, it was again an impulse coming from logicians—not from psychologists—that put logic back in the psychological ballpark. In the rich intellectual debate generated by the problem of the foundations of mathematics, the formalist school has a special role for our story.

[9]See Skinner (1957), pp. 437 ff.

Let us distinguish Hilbert's specific program for the foundation of mathematics from the general formalist approach. The former required that a consistency proof for arithmetic be obtained with finitary means, and Gödel's second theorem killed it by showing that this cannot be done. The latter more general framework, however, according to which logic was conceived as the science of signs and their transformations, is not necessarily tied to Hilbert's foundational program. Although this conception falls short of explaining mathematics, it might well be sufficient to explain thinking. It is interesting, and often overlooked, that Hilbert had already realized that his conception would open important new perspectives not only for mathematics, but also (and especially) for the relation between logic and reasoning. Thus when rejecting Brouwer's scorn at this "material" aspect of logic, so cherished by formalists, Hilbert (1927) wrote:

> The formula game that Brouwer so deprecates has, besides its mathematical value, an important general philosophical significance. For this formula game is carried out according to certain definite rules, in which the *technique of our thinking* [his italics] is expressed. These rules form a closed system that can be discovered and definitively stated. *The fundamental idea of my proof theory is none other than to describe the activity of our understanding to make a protocol of the rules according to which our thinking actually proceeds* [italics added]. Thinking, it so happens, parallels speaking and writing: we form statements and place them one behind another. (p. 475)

To this idea of proof theory as a description of the activity of the understanding, Hilbert added a particular interpretation of Kant's thesis that something must be presented in intuition in order for thought to produce knowledge. In the case of mathematics, the extralogical intuitive necessary element is *the symbol itself*, a concrete object which, just by virtue of its *shape*, can trigger elementary processes of recognition, matching, concatenation, and deletion and allow a rule to be applied:

> Kant already taught . . . that mathematics has at its disposal a content secured independently of all logic and hence can never be provided with a foundation by means of logic alone; . . . Rather, as a condition for the use of logical inferences and the performance of logical operations, something must already be given to our faculty of representation, certain extralogical concrete objects that are intuitively present as immediate experience prior to all thought. If logical inference is to be reliable, it must be possible to survey these objects completely in all their parts, and the fact that they occur, that they differ from one another, and that they follow each other, or are concatenated, is immediately given intuitively, together with the objects, as something that neither can be reduced to anything else nor requires reduction. This is the basic philosophical position that I consider requisite for mathematics and, in general, for all scientific thinking, understanding, and communication. And in mathematics, in particular, what we

> consider is the concrete signs themselves, whose shape, according to the conception we have adopted, is immediately clear and recognizable. (Hilbert, 1925/1967, p. 376)

Liberated from its interpretive parts, the passage contains the clear suggestion that logical rules of thought can be seen as procedures activated only in virtue of the *form* of the signs with which they are formulated. In our terms, we may say that a thinking process can be seen as a set of transformations on symbols according to rules exploiting a high-level concrete property of the symbols—their shape.

Hilbert's two great intuitions—that a thought process is a kind of proof, and that a proof is a transformation of symbols according to rules acting on their form—gave formalism a completely different turn. The syntactic nature of logical rules, together with the statement that our thinking is nothing but the application of rules in a certain order, was opening again the way for undertaking an investigation of the logic of thinking, but this time with a clearer understanding of the nature of the rules and of how properties of their forms could be used to generate deductions.

Thus formalism provided two essential ingredients for the mental-logic cake. Others, however, were still missing. What logic needed to be ready for psychological investigation was, on the one hand, a more intuitive presentation of formal systems and, on the other hand, a model of how a physical structure could use a formal system to carry out derivations. Gentzen provided the first ingredient, and the second one emerges from Turing's work.

KNEAD FORMS UNTIL THEY FEEL NATURAL TO THE TOUCH: GENTZEN AND NATURAL LOGIC

Hilbert's axiomatization of logic was meant to capture the set of theorems by using few initial sentences and the smallest possible number of inference rules. This system, however useful for handling many domains of logic and mathematics, fails to be appealing even to represent the particular kind of reasoning it aimed to represent, namely, logical and mathematical reasoning. Far less could it be proposed as a model of how human reasoning generally proceeds. Gentzen's natural deduction took care of this problem. An alternative treatment was developed for some parts of logic, in which axioms disappeared, leaving many inference rules in their places. Gentzen (1964) consciously intended to provide a system of deduction "as close as possible to actual reasoning":

> The formalization of logical deduction, especially as it has been developed by Frege, Russell, and Hilbert, is rather far removed from the forms of deduction used in practice in mathematical proofs. Considerable formal advantages are achieved in return. In contrast, I intended first to set up a formal system which comes as close as possible to actual reasoning. The result was a "calculus of natural deduction." (p. 68)

He noticed that in actual mathematical reasoning, appeal to formal axioms is rare, whereas many—but not too many—forms of inferences are used. He set himself the task of specifying those inferences. To this purpose, he introduced different kinds of calculi in which each logical symbol occurred in two rules specifying when a formula containing it could be introduced and when it could be eliminated in favor of other formulas not containing that symbol.[10] This central difference from axiomatic systems makes proving a theorem a completely different, more natural activity than proving it starting from axioms. Yet, nothing is lost in natural deduction: The crucial metatheoretical properties of axiomatic systems are all preserved.

Another difference with Hilbert-style logics is worth noticing. Ideally, an axiomatic system aims at maximal economy, both in axioms as well as in rules. Gentzen, instead, provided rules for all the logical connectives, even if they were logically redundant. There are technical reasons for it, but there is also an important philosophical motivation. Gentzen intended to provide a system as close as possible to actual mathematical reasoning, and from this point of view economy in rules could not be a value in itself. If mathematical natural reasoning seems to require even redundant rules, so be it. A related point has to do with the interdefinability of connectives. In classical logic equivalence laws, such as, for example, $(A \rightarrow B) \leftrightarrow (\neg A \vee B)$, render connectives interdefinable and thereby dispensable in principle. However, this goes flat in the face of another natural intuition, the feeling that each connective has a separate meaning. Natural deduction also provided a way to account also for this intuition. Even if one adopts classical rules for connectives, and thus the equivalence laws are still theorems, from Gentzen's point of view such extensional equivalences do not show that one can dispense with the rule. If we use them when following a path of thought, if they appear to govern the functioning of certain special words lexicalized even in normal natural language, such as *and*, *or*, *if*, or *all*, then they are likely to correspond to separate units of meaning and they deserve a place in a system of mathematical reasoning.

---

[10]With the exception of classical negation. In the intuitionistic case, negation is governed by the elimination rule $\frac{A, \neg A}{\bot}$ and by the introduction rule $\frac{\begin{array}{c}A\\ \vdots \\ \bot\end{array}}{\neg A}$, whereas to get classical logic a special extra elimination rule expressing double-negation redundancy is needed: $\frac{\neg\neg A}{A}$. This rule for Gentzen spurs the "harmony" of a system in which to each connective just one elimination and one introduction rule correspond.

In short, Gentzen's natural deduction achieved for mathematics what Gentzen wanted to achieve: It showed how a proof can be at the same time formal and natural. The route from formal logic to natural reasoning was opened. Gentzen did not provide a *psychological* system of rules for reasoning—he could not, because the psychology to go any further was not yet available—but because *mutatis mutandis* his arguments hold for reasoning in general, he contributed another basic ingredient to its development. He shaped the form of the rules and showed how formal logic (this time, a real discipline separated from psychology) and intuitive logic may be not so distant from each other, after all.

WHIP COMPUTATIONS UNTIL THEY ACQUIRE THE RIGHT CONSISTENCY: THE COMPUTER METAPHOR AND FUNCTIONALISM

The other necessary step for transforming mental logic into a psychological thesis was to develop fully Hilbert's intuitions about how rules exploit the "concrete" aspects of signs. Notoriously, it was Turing who offered the abstract model of a shape-driven machine, which showed how a physical mechanism could perform operations once considered to be typically mental. Turing's intuitions on the role and possibilities of his new machines were of exceptional importance. They finally led to the thesis that the mind can be seen as an information processor, which in turn gave substance to the idea of a mental logic. However, in the development of this view two stages should be distinguished. The first stage, explicitly endorsed by Turing, is expressed by the thesis that *a computer can be a mind*, namely, that certain kinds of properties once attributable only to humans can also be appropriately predicated of other physical configurations. This thesis, however revolutionary, leaves completely undetermined the nature of the operations of the mind. It may imply that mental processes can be *simulated* by a machine but leaves it open the possibility that the mechanisms and procedures intervening in the two physical beings—the *simulandum* and the *simulans*—are totally different. All that the thesis requires is that the input/output relations typical of mental processes are preserved by the simulation device, and simulation only establishes an extensional identity between two physical configurations, but leaves it undetermined whether these share the same *psychology*.

Turing's well-known test for intelligence, entirely based on the control of input/output relations (at least prima facie), uses only dispositional language and is thus compatible with conceptions that do without mental vocabulary, such as behaviorism or logical behaviorism. It does not *require* laws mentioning internal states, and if it does, it doesn't require that they

apply to *both* minds *and* computers. To see this, imagine that a computer passes Turing's test; it may still be the case that when we reason we exploit lawfully connected inner states, but the computer does not. Suppose, for example, that *we* parse the computer's responses to our questions by using a grammar for a language. Then, a part of our psychology would be described by linguistic theory. But when the computer "reads" our questions, it consults a huge database of predefined forms and outputs answers to each question that we plausibly may ask by mere physical shape-matching, without really parsing our input.[11] In this case, although the computer may happen to pass Turing's test, we and it would not fall under the same psychological laws: Linguistics would be true for us, and false for the computer.

Turing did intend to go beyond this minimal claim; witness the fact that he proposed tests for machine intelligence couched in more "cognitive" vocabulary.[12] However, his main concern was not to dispel the prejudice that *we* cannot be like machines, but rather to reject the reverse thesis that *machines* cannot be like us. For him, the question of whether the inner structure of mental processes and the computational states of machines are captured by a common set of laws was not at the forefront. He did use analogies to mental processes for finding algorithms, but this was meant to be a heuristic strategy based on introspection rather than a research program.[13] His caution also is shown by his willingness to explore all possible sorts of machines, whether "classical," as we would call them, or "connectionist," with the main task of finding agreement in behavior between minds and machines, rather than agreement in laws.

Turing's ideas were of exceptional value, but for various reasons—mostly due to the widespread behaviorist attitude in psychology and philosophy—further time had to pass before they could be elaborated into a fully coherent program. As Hegel wanted it, philosophy always comes at dawn. The new paradigm had to wait for the development of the theory of computable functions and for the first successes of artificial intelligence before being taken seriously by philosophers. The big further step—the reverse thesis that *we* are like computers—came with function-

[11]The computer could not answer *all* questions we may possibly ask, because it wouldn't include a productive language, but lack of productivity would not forbid it to pass Turing's test. The test only requires that we could not *plausibly* tell a computer from a human, not that we could not *possibly* do it.

[12]Specifically, he proposed that a machine that learns (i.e., that modifies its instruction) should be considered intelligent; see Turing (1947b, pp. 122–123).

[13]For example, when outlining a program for playing chess, he wrote, "If I were to sum up the weakness of the above system in a few words, I would describe it as a caricature of my own play. It was in fact based on an introspective analysis of my thought processes when playing, with considerable simplifications. It makes oversights which are very similar to those which I make myself" (Turing, 1947b, p. 294).

alism. Functionalism explicitly defended the thesis that the psychological vocabulary *is* computational vocabulary, and that the natural kinds described by psychology are not organisms but computational devices. Fodor (1981) presented it as a *fairly recent* contender in the arena of the theories of mind:

> The real point is that, if we want a science of mental phenomena at all, we are required to so identify mental properties that the kinds they subsume are natural from the point of view of psychological theory construction. . . . Now, there is a level of abstraction at which the generalizations of psychology are most naturally pitched and, *as things appear to be turning out*, that level of abstraction collapses across the differences between physically quite different kinds of systems. . . . [I]f we wanted to restrict the domains of our psychological theories to just us, we would have to do so by ad hoc conditions upon their generalizations. Whereas, what does seem to provide a natural domain for psychological theorizing, at least in cognitive psychology, is something like the set of (real and possible) information processing systems. (pp. 8–9, italics added)

We are now well beyond the 1960s. Only then were logic and philosophy ready to nurture the development of mental logic as a serious psychological hypothesis. Everything was there, except the psychology.

MIX WITH EMPIRICAL PSYCHOLOGY: ALMOST GETTING THERE

Ideas must be ripe before becoming productive. I argued that the although the idea of a mental logic has been around for centuries, it could not develop because the ingredients necessary to give substance to it either were not available or were seen from a wrong perspective. When, finally, all the conceptual ingredients did become available, the psychology for testing mental-logic theories was still missing. Another 20 years or so had to pass before experimental techniques were sufficiently developed to begin asking nature the right questions in the right way. It is true that Piaget and his school (e.g., Piaget, 1953) spoke a language externally akin to mental logic, but in their hands the thesis was never even spelled out clearly enough to see what it said or how it differed from other alternatives (see Bonatti, 1994; Braine & Rumain, 1983). If one eliminates Piagetianism, which did more harm than good to the mental-logic hypothesis, only a handful of articles and proposals remain that can be considered real attempts to transform it into a testable psychological theory, and until very recent years most of them fell short of being satisfactory, for one reason or another.

Henle (1962) was probably the first attempt to test experimentally a mental-logic hypothesis, but the limits of her work were important. She took natural logic to be classical logic—not even in natural deduction form—but she gave no argument for this assumption. Her method was entirely based on direct analysis of subjects' explicit conscious justifications, but those may not be a good indicator of the real processes involved in reasoning. She assumed that the syllogism has a special role in reasoning, but this too was an unjustified assumption. In short, the merit of Henle's study was less in her arguments and methodologies than in the theses she presented.

It was only in the 1970s that the mental-logic hypothesis marked significant improvement towards its transformation into a psychological theory. In those years natural deduction became the standard form of the theory, and purely qualitative analysis was abandoned in favor of more quantitative methods often accompanied by computational implementations. Nevertheless, the proposed theories still lacked a sufficient psychological motivation. Osherson (1975a) and Johnson-Laird (1975) proposed models of rules that shared many common points and diverged for computational options. Whereas Osherson's model was conclusion sensitive and applied its rules in strict order trying to minimize the difference between premises and conclusions, Johnson-Laird's (1975) program blindly generated consequences of the premises, regardless of the form of the conclusion. Both systems included good and new ideas, but, interestingly for us, both were clearly more concerned with the algorithmic implementation of rules than with their psychological plausibility. For example, in both systems the basic rules were selected somewhat arbitrarily. Both systems took a one-way approach to the role of a conclusion in a problem, one being conclusion sensitive and the other conclusion blind, but neither provided evidence that subjects exclusively or preferentially follow one of those strategies. Likewise, although Johnson-Laird's system introduced the distinction between primary and secondary inferences that remains central also in the system presented in this book, his way to sort the rules in the two classes was not psychologically motivated. It is revealing that at the end of his 1975 article, Johnson-Laird admitted to not being sure of how to test his model empirically, and he concluded that this was a really difficult, almost impossible, task. Clearly, psychological reality was not at the forefront.

In fact, both Osherson and Johnson-Laird could hardly have done any better. They had run into another tricky problem. They lacked another crucial ingredient for a mental logic, a really difficult one to find and to calibrate: the right balance between algorithmic implementation and psychological justification for a mental logic system. Only recently (e.g., Braine, 1978; Rips, 1983) did some researchers start tinkering with it, to find out the right proportions between the two ingredients. This work is still at its beginnings, and much is still needed to nail down the rules for reasoning, their implementations, and the flow of information involved in a reasoning process.

HAVE A TASTE OF THE CAKE: GETTING THERE

So developing a mental-logic theory was not so easy after all. Many ingredients needed to be found and their proportions carefully weighed before a good mixture could be formed. Logic had to be separated from psychology, but this only happened after Boole, with the development of modern logic. Its formal nature had to be put in focus, and the logical properties of forms had to be discovered, but this only happened when logicians started working on the foundations of mathematics, at the beginning of this century and well into the 1930s. The connection between forms and machines acting on forms had to be seen, but this was clear only with Turing and his notion of computability. The philosophical consequences of the new view about mental processes implicit in Turing's ideas had to be drawn, but this only happened some decades later, when philosophical functionalism ripened. Only then could psychology revive the mental-logic hypothesis and look at it with a clearer vision of the claims being made. Psychology too had to make progress. It had to abandon the ancient commonplace that the laws of reasoning can be discovered by immediate introspection. Also, this step was slow to come. It had to devise experimental techniques to investigate the rules, and to test the psychological reality of the proposed algorithms by implementing them in the conditions in which they can be deployed. All of it is very recent development, and much of it has yet to be written.

It should no longer be a surprise that mental logic started developing fairly recently. In fact, it has a development as recent as that of its supposed sexier alternatives, such as models, interconnected networks, space phases, or what have you. The reason is simple: They all require pretty much the same conceptual ingredients to be developed, and they all require a well-developed empirical psychology. There is no dead research program in the area of reasoning; there are only theories in better or worse shape.

"OK, fine. But even if the long history of mental logic does not count against it, its *present state* does! Even now, with all the ingredients available, the mental-logic dough has not leavened. Compare with X [substitute X for models, interconnected networks, space phases, or what have you]: X has gone so much further!" Again, there is an answer. This time, the answer is "No and No." No, it is false that there are better alternatives to mental logic, and, No, it is false that mental logic has not begun to be productive. Have a taste of the cake by reading the rest of this book.