Is Thought (Il)logical or Logic (Un)thinkable?*

An Exploration of the Relations between Thought and Logic

Luís M. Augusto

MIag, Institute of Philosophy, University of Porto (postdoctoral researcher) / FCT (postdoctoral fellow).

Abstract

The issue of the (il)logicality of our thought processes is highly pertinent, as it commonly is a factor in the diagnosis of thought disorders present in, for instance, schizophrenia. But, given the contemporary plethora of extensions to and rivals of classical logic, what is to be considered 'logical' today? On the other hand, classical logic claims to describe the way we actually think, but, aiming at an a priori description – or even regulation – of human thought processes, its rules often appear to be simply unthinkable. We thus need to establish an intersection where thought is logical and logic is thinkable. This paper analyses this intersection in the light of the 'new' logics and of the claims of the 'old', classical, logic confronted with actually observed thought processes, normal or clinically seen as abnormal.

Keywords: Logic(s); Thinking; (Il)Logicality; (Un)Thinkability

I. Introduction

1. A Minefield

The relations between thought and logic are far from well understood, and the interested parts take sides with the radicalism that often characterizes only partially or insufficiently mastered scientific matters. But contrary to more typical problem fields, where usually one debate opposes two confronting factions, these relations have given origin to a whole minefield of actually *several* debates and controversies. On one parcel of the minefield, logicists oppose defenders of (logical) psychologism, these claiming that logic depends on thought, that is, logical laws are psychological laws (Mill, 1843; Erdmann, 1892), and the former arguing that logic is independent, wholly a priori in relation to thought, being as a matter of fact the body of normative, objective rules from which human thought draws – or should draw (Boole, 1854/1958; Frege, 1884/1953; Hilbert, 1931/1996; Husserl, 1900/1970).[1]

*Status of the paper as of Nov. 24, 2011: Submitted.

[1] The cited texts reflect at the same time the highly localized source of the dispute and its largely intractable character, with not much (if any) progress having been achieved since its inception. For an analysis of the problem, see Kusch (1995).

In another, more contemporary region of the minefield, computationalists, heavily relying on some form of logic, claim that thought is a form of computation involving symbols, concepts, and/or representations in general (Fodor, 1975; Fodor & Pylyshyn, 1988; Neisser, 1967), and by this they oppose the connectionists, who focus their attention on interconnected networks of more or less basic units of processing (e.g., neurons) and often claim to dispense with symbolic representations and the associated logical(-like) rules of computation (Rumelhart & McClelland, 1986; Elman, 1991). More recently, an attempt has been proposed to fit logic into a more constructivist theory of thought and its relations to logic named logical cognitivism by its conceiver (Hanna, 2006). In yet another portion of the mined terrain, focusing on the psychology of reasoning, several recent approaches simply reject that logic might play any relevant role in thought (e.g., Evans, 2002; Gigerenzer, 2000), while efforts are being made to elaborate (on) mental logics (e.g., Braine & O'Brien, 1998) or, emphasizing content but still having to do with (some form of) logic, to devise mental models (e.g., Johnson-Laird, 1983) or specialized inferential mechanisms shaped by evolution (e.g., Fiddick, Cosmides, & Tooby, 2000).

This profusion of debates and controversies shows how central the question is in contemporary research, a centrality that is indeed justified by the benefits we might obtain from a more accurate understanding of this issue. For instance, research on severely impairing mental 'conditions' involving psychosis or simply psychotic symptoms[2] (paradigmatic case: schizophrenia) has hardly improved our understanding of the condition, from its dubbing in the 1840s to the vacillating contemporary use of the term in the current versions of both the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR) and the *International Statistical Classification of Diseases* (ICD-10; see Bürgy, 2008), to a great extent precisely because we lack a comprehension of the relations between thought and logic: are psychotic thought processes plain illogical (i.e. contrary to logic), alogical (beyond, without logic), or do they follow some sort of non-classical, perhaps deviant logic? The fact is that, though unable to exactly pinpoint the problem, the clinicians responsible for diagnoses often intuit that, from the viewpoint of logic, something is awry (e.g., Akiskal, 2008). These relations between thought and logic are, in turn, apparently interconnected with language disorders (McKenna & Oh, 2005), but we have failed to distinguish the contributions of language and 'logic&language' to thought disorder. Thus, though thought disorder is often a factor in the diagnosis of psychotic conditions (see McKenna & Oh, 2005, Chapter 2), those responsible for this diagnosing will more often than not show great difficulty, if not embarrassment, in explaining what exactly thought disorder is and how it is supposed to help in the diagnosis of a psychotic disorder such as schizophrenia. This, in turn, causes the relegation of these and other 'conditions' to the field of psychiatry, where commonly symptoms matter even in the lack of their understanding, and their suppression by pharmacological means alone is seen as a successful therapy.

But the profusion of antagonisms regarding the relations between thought and logic may to a great extent be due to the lack of a precise understanding of what it is we talk about when we talk about thought, logic, and thought *and* logic. By *thought*, one may mean something as broad as *cognition*, or something as restricted as *reasoning*. *Logic*, in turn, also needs a reappraisal in face of a growing – and seemingly unstoppable – plethora of so-called non-classical logic systems that are

[2] Typically, these comprise delusions and hallucinations, as well as speech, thought, and/or personality disorders, which, individually or concurrently, lead into what can be seen as a loss of touch with reality.

supposed to extend or even replace classical logic, the *one* logic commonly appealed to in the relations between thought and logic.

This apparently solely academic issue is of the greatest 'practical' importance. For instance, for the psychiatrist – and often for the psychologist, who more often than not also resorts to the DSMs and ICDs at hand, delusions are a symptom of psychosis (or of a psychotic disorder, such as schizophrenia). Delusions are seen by clinicians as strong beliefs whose content is impossible (e.g., I'm Napoleon Bonaparte. Jesus tried to shoot me. I'm dead.) because it is simply false. This entails the explicit or implicit acceptance by the clinician of the bivalent, classical logic that separates propositions into either true or false. The person who believes a falsity and rejects any evidence against it, powerful – that is, true – as it might be to the eyes of the 'approved' elements of her/his community, is said to be delusional because s/he is illogical. Take the individual who believes s/he is dead, and persists in this belief even when agreeing that dead people do not talk, i.e., that only people who are not dead can talk: this person is clearly thinking against the pillars of bivalent logic, the principle of contradiction, according to which a proposition P and its negation, representable as $\neg P$ (also: $\sim P$), cannot both be true, and the law of excluded middle, stating that either P or $\neg P$ must be true. Yet, according to multi-valued logics, which accept three or more truth values, a proposition like "I am dead" might be true, false, or undefined if put in the mouth of, say, Napoleon Bonaparte, Barack Obama, and Daffy Duck, respectively. In the same way, the proposition "I'm Napoleon Bonaparte" might be true (if uttered or believed by Napoleon Bonaparte himself), false (if, for instance, uttered or believed by the cartoon character Daffy Duck qua Daffy Duck), or simply more or less likely (if uttered or believed by someone who looks very much like Napoleon; has the same ideas as Napoleon; claims to be a reincarnation of Napoleon; etc.). What are you to do if you are a 'multi-valued clinician'?

2. The Task Ahead

Thus, though much work has been done on the connections between thought and – mainly classical – logic, it might help to go back to basics and try to determine, first, (a) *what exactly thought processes are* and (b) *what we (still) mean by logic today*, then see (c) whether logic has anything to do with thought after all, that is, *whether logic is (un)thinkable*, and finally settle (d) *when and in which way we can say that thought is (il)logical*. Question (a) has already been addressed elsewhere (Augusto, 2011), and here I shall focus on – re(de)fining the concept of – logic and the relations between it and thought. This can be done while remaining wholly agnostic regarding the disputes that oppose computationalism to connectionism; this debate seems misguided in essence, as, clearly, the positions involved are not necessarily antagonistic, but rather complementary: given that human thought takes place in the brain (we have no evidence of thought in cases of brain death, for instance), it is highly plausible that symbolic manipulations – taken here to be the essence of thought – are carried out by interconnected networks of, say, neurons. Regarding the controversies within the more restricted field of psychology of reasoning, though they respect only a (small?) part of thought, especially deductive thought, they are seen as informative for a more encompassing theory of thought, but I shall skip them entirely; in this text, mental logic, among the contenders, will be the only one approached.

Given the intractable, chicken-or-egg character of the debate involving logicism and psychologism, I would rather remain agnostic regarding it. This, unfortunately, can perhaps not be: from the moment one so much as accepts the possibility of non-classical logics, one is taking the world at large, and psychological factors in particular, into account in logic issues; more specifically, logic can no longer be seen as a kit of a priori, objective rules that dictate the way we (should) think, but only as a (largely?) a posteriori rationalization of the way we actually think, which, as shall become evident below, is part of our empirical nature. This, in turn, entails, I suspect, an intimate connection between language and logic. But this view does not in principle force us to choose sides in the mentioned debate: in accepting deviant logics as logics de facto, the normative notion of *logical law* becomes rather opaque, hindering at the same time the claim that logical laws are psychological laws (e.g., do infinite valued-logics really reflect the way we actually think?) and the aspiration that thought processes should conform to a priori, objective logical laws and rules. My starting point is that thought and logic, though connected in an obvious way, can be independent in the sense that thought can be illogical or alogical without for that being less thought, and logic can be thinkable or unthinkable without for that being less logic. Plainly, there are logical laws that no one honestly expects to be descriptive of particular human thought processes, let alone that they should actually be normative, and there are thought processes, in particular creative and/or heuristic processes, that appear to defy any logic. The challenge is to find the intersection between thought and logic that allows us to talk of (il)logical thought and/or of (un)thinkable logic (see Fig. 1).

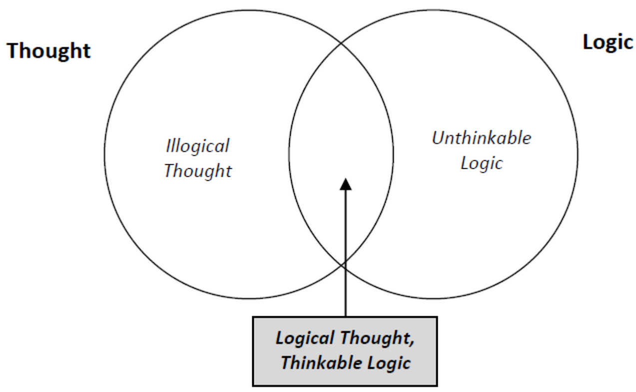

Fig. 1: The relations between thought and logic.

The main objective of this paper is to provide a working basis from which to develop theories on the relations between thought and logic, which, in turn, might contribute to a better understanding of such vital concepts as *(ir)rationality*, *mental (in)sanity*, *mental (ab)normalcy*, etc. Although the logician may learn from this text, its primarily aimed at audience is the researcher of the human mind: the cognitive scientist, in general, and the clinician in particular. With this aim and audience in view, elementary logic issues will be approached in more detail than that required by logicians or logic-trained readers, without for that entering into technicalities; that is, a broad, basic overview of classical and non-classical logics will be provided. Regarding the important issue of the connections between logic and language, I will not approach it in this text. However, while not claiming that they are reducible one to the other, I take them to be intimately connected, and this view may sometimes surface in this text.

Needless to say, I think that the questions above, in particular questions (c) and especially (d), are relevant. As said, much work has been done regarding the relations between thought and classical logic, but the non-classical logical systems have been all but neglected in the debates around these relations; this paper innovates in bringing them into the debate.

Finally, I am as far as could be from wanting to see any of my answers applied as cutting devices between rationality and irrationality, sanity and insanity, normalcy and abnormalcy, etc. This caveat risks weakening the theses here advanced, but that is a necessary risk: in matters of this kind, where *in the end* ethical issues are concerned, theories should be taken with a healthy degree of suspicion rather than being unwarily embraced. Allying cognitive science with hermeneutics, the aim here is to help to develop our understanding of these issues without for that pretending to explain reality *ipsis verbis*.

II. Thought & Logic

1. Refining the Concept *Thought*: A Summary

As indicated above, I have approached the rethinking of the concept of thought in Augusto (2011), and here I will provide only a summary of the main conclusions reached in that text. To begin with, thought is all about *representations*. That is, we think with 'mental stuff' that results partly or entirely from the irritation of our sense organs and the consequent activation of specific areas of our brains. Whether this 'mental stuff' that I call *representations* is symbolic in nature, image-like, ideational, conceptual, etc, and whether you choose to call it ideas, proximal stimuli, abstractions, concepts, etc, is at this point of our scientific knowledge impossible to determine and thus best left unspecified: representations, the 'mental stuff' we think with, is simply to be understood as the stuff with which we perceive and manipulate the world, and in this text any of the above might appear in this text as a synonym for or token of *representations*.

So, this is the 'stuff' of thinking. Now, for the way we do it: we think by creating, categorizing, associating, and combining in other ways this 'mental stuff'. Although thought processes may involve less than these (hypothetically in severe cases of hydroanencephalia, for instance), or more than these (an unlikely hypothesis, as in a divine being), we can see them as the *basic* processes involved in thinking: stimuli from the world at large (which includes us and in which we are included) can cause at any instant one of these processes to occur in isolation or together with any of or all the others. They are not wholly discrete, but rather continuous processes; nevertheless, it would be a mistake to emphasize this processual continuity. For instance, forming a representation implies its categorization (we do not form wholly non-categorisable representations), but one can categorize a representation without having the need to create it: one can simply retrieve it from memory. Then, when retrieving it from memory, one is inevitably associating it with other representations from the same or different categories; one may also recategorize a representation, which will entail a certain recreating; etc. But the main point is that by and large the list above is a hierarchy, in that each upper level is dependent on all the lower levels that precede it. That is to say that while categorization of representations is only dependent on their formation or creation processes, combining them implies their formation, categorization, and association.

This will do, but will also be elaborated upon, in what follows.

2. Re(de)fining *Logic*

2.1. Classical Logic and 'Classical' Set Theory

In general terms, logic is the systematic study of correct thought or reasoning. More specifically, logic is the science of *deductive* reasoning, or valid inference, that is, it is the science of which arguments are valid. For instance, the following deductive argument, composed of two premises and one conclusion, is valid:

All men are mortal.
Socrates is a man.
$\therefore$ (= Therefore) Socrates is mortal.

It indeed makes sense to infer that if all men are mortal, and Socrates is a man, then Socrates must be mortal. That is to say that we cannot accept that the premises are true and yet affirm that the conclusion is false. This is the definition of *valid* argument or inference: its conclusion cannot be false so long as its premises are true.[3] So far so good, but then, what a true premise is, is far from a simple matter. Moreover, it actually is not a particular argument that is valid, but its *argument form*: logic is not interested in individual arguments, but in their *logical form*. Let the above argument be 'translated' into a more logical language:

If an entity is a man, then it is mortal.
Socrates is a man.
$\therefore$ Socrates is mortal.

This is said to be a more logical language because it follows a strict syntax, in this case comprising the operator *if ... then* ($\rightarrow$, in logical notation), known as material implication or material conditional, joining two propositions, "x is a man" (representable as, e.g., P) and "x is mortal" (Q), whose subject, x, is undetermined.[4] We say that the argument above is valid because any argument with its form,

[3] Logic also studies inductive reasoning, but there is no talk of *validity* regarding inductive arguments; these can, at best, be 'strong' or 'weak,' strength being the property that determines their probability. Let us take an example of an inductive argument:

Every n observed crows are black.
$n + 1$ is a [non-observed] crow.
$\therefore n + 1$ is black.

Obviously, though the premises in this argument are true, the conclusion need not be so, as it is possible (though perhaps not probable) that there is a non-observed crow that is not black. This reduces greatly (or even entirely) the normative character of logic, which makes inductive logic of no import for our ends in this particular study. This, however, does not mean that inductive logic is not relevant for a cognitive approach of human thinking processes.

[4] For the initiated, a more immediate translation would be of the form "All *A*s are *B*s", but I am leaving this for the introduction of set theory below. Nevertheless, we can already easily see that the argument above is valid, because it is of the form "All *A*s are *B*s. *X* is an *A*. $\therefore$ *X* is a *B*."

$P \rightarrow Q$
P
$\therefore Q$

a form known as *modus ponens*, is necessarily valid. This means that an argument like

If you deedle-dee, then you booble-boo.
You deedle-dee.
$\therefore$ You booble-boo.

is also necessarily valid, though it is simply impossible to decide whether the premises are true or false,[5] the two truth-values allowed by the semantics, or truth-functionality of this logic. In technical language, this means that the logical definition of validity is *monotonic*.[6] Given this disregard of content in face of the importance of form alone, allied to a clear-cut, absolute distinction between truth and falsity – known as *bivalence* –, this science of valid inference founds itself on the already mentioned laws of contradiction, representable as $\neg(P \wedge \neg P)$, and of excluded middle, $P \vee \neg P$ (where '$\neg$' is the logical unary operator of negation, '$\wedge$' is the binary operator known as conjunction, and '$\vee$' is the binary operator known as disjunction), which, in turn, force the following truth-functions of logical expressions, known as truth-tables, where 0 stands for false and 1 for true (Figure 2).

P	¬P
1	0
0	1

Negation (NOT)

P	Q	P ∧ Q	P ∨ Q	P → Q	P ↔ Q
1	1	1	1	1	1
1	0	0	1	0	0
0	1	0	1	1	0
0	0	0	0	1	1
		Conjunction (AND)	Disjunction (OR)	Conditional (IF, THEN)	Bi-Conditional (IF AND ONLY IF)

Fig. 2: Truth-tables for classical logic.

The above are the basics of what is today known as classical logic, the adjective *classical* conveying the fact that it was created by Aristotle and survived until well into the 19th century without major changes or progress; the mathematical-like formalism and symbolization that we briefly approached are later developments introduced by the *mathematization* of logic, which was more precisely a retake of logic, after more than four centuries of neglect, to apply it in the service of

[5] That is to say that it is impossible to decide whether the argument is *sound*.

[6] As a matter of fact, *monotonicity* regarding logical validity means that new information cannot 'undo' a deductively valid inference (more formally, if a sentence can be inferred from a set of premises Γ, it can be inferred from a set of premises Δ that extends Γ). For our purposes, however, it will suffice to say that the validity of an argument form is not affected by whatever information we might have regarding the premises; for instance, the information that *deedle-dee* and *booble-boo* are utterly meaningless terms (i.e., they have no reference), an information that can be considered as a supplementary premise, does not intrude in the validity of the argument. This is actually an important feature as, as we shall see, one of the motivations of non-classical logical systems was precisely the need to represent reasoning that takes new information into account (defeasible reasoning).

mathematical proof.[7] This mathematization can be said to have started with Boole (1847/1948), and with Frege (1884/1953) it turned into logicism, firstly the stance that mathematics is identical with logic, and, derivatively, that (mathematical) thought is governed by logical rules.

This itsy-bitsy piece of history of logic interests us,[8] because, as we shall see in the next section, it was precisely this mathematization that eventually opened the Pandora's box leading to the proliferation of logic systems claiming to be candidates to *replacing* classical logic. But before advancing to this dramatic development of logic, it is essential to mention that this mathematization also motivated the birth of set theory, whose main characteristic of interest for us is its 'bivalence': an object either is a member of a set, or it is not (that is, a set is a well-defined collection of objects). For instance, the number 2 is a member of the set of the positive integers, but it is not a member of the set of the negative integers. In the mentioned logicization of mathematics, set theory established binary relations (e.g., membership, inclusion, identity) and operations (e.g., union, intersection, complement) motivated by the need to define formally all mathematical concepts, but it is obvious that set theory can be applied to any *universe* of objects, as the example below of the universe of living organisms shows.

Let us take the argument introduced above: the proposition "All men are mortal" simply means, in terms of set theory, that all the members that belong to the set of men (let it be the set A with a_m members, where m is a finite number) also belong to the set of mortal beings (set B); formally, $a_1, a_{2,}a_3, ..., a_{1,000,000,000,000}, ..., a_m \in B$ (all a belong to B).[9] Several other set-theoretical relations and operations can express the same proposition "All men are mortal":

$A \subset B$ (the set A is contained in the set B)

$A \cap B = A$ (the intersection of the sets A and B is the set A)

$A \cup B = B$ (the union of the sets A and B is the set B)

$A \cap B^c = \emptyset$ (the intersection of the set A with the complement of the set B is the empty set)[10]

Set theory, especially *naïve* set theory, so called because it uses natural language to define sets (that is, it defines sets informally),[11] using Euler and Venn diagrams to pictorially represent set-theoretical relations and operations (see Figures 3.1 and 3.2),[12] can thus be seen as a description of the correct ways we think about sets.

[7] Reason why it is also a *logicization of mathematics*, namely in the face of doubts concerning the reliability of spatial intuition in mathematical reasoning: mathematics, as a rigorous science, was thought to require another foundation than intuition, and logic first appeared as that foundation.

[8] The reader interested in details of this long history should consult Kneale and Kneale (1962); the reader should be aware that this is a text for the initiated.

[9] More correctly, $\forall a : a \in B$ (where '$\forall$' is read 'for all'), or, more completely, $\forall a \in A : a \in B$.

[10] That is, there are no non-mortal men. Formally, the complement of a set A, denoted by A^c, is the set whose elements belong to the universe ($\mathbf{U}$) but do not belong to A. In this case, the universe consists of all the living organisms. It is easy to see that in this case B^c is itself an empty set, i.e., there are no non-mortal living organisms.

[11] In technical mathematical jargon, this means non-axiomatically.

[12] Contemporarily, only Venn diagrams are commonly used, as they represent all the relations and operations that Euler diagrams can represent, while actually extending the representative possibilities. Euler diagrams differ from Venn diagrams in that in these the universe is taken into consideration, being represented by an enclosed rectangle in the plane; moreover, in Venn diagrams intersections of curve interiors can be empty. In a Venn diagram, typically three circles represent

This 'correctness' is strengthened if one realizes that set theory is, in an important way, logic in disguise,[13] to the point that the former can be simply labeled 'logic of classes,' namely when the algebra of sets is meant (see Appendix, Fig. 12). This is disputable, and arguably attributable to an overestimation of the kinship between the set-theoretical notion of membership and the logical concept of predication, which are, in turn, connected to the intermediate notion of attribution of attributes[14] (Quine, 1970, p. 66), but at the very least it is problematic to distinguish set theory and logic. Not the least because the logical principles of non-contradiction and of excluded middle, dependent on logical bivalence, fit set theory like a glove: an object x cannot belong and at the same time not belong to a set P, that is, an object x either has property P or it does not.[15] But regardless of the opinion of logicians and mathematicians on the matter, this 'similarity' or suspected identity between logic and set theory is of extreme relevance to us, as we shall see; for our ends, I shall see set theory as, if not 'logic in disguise,' a part of logic.

2.2. Non-Classical Logics and 'Non-Classical' Set Theory

As said above, the mathematization of logic opened the Pandora's box that would throw doubt on the status of classical logic: it motivated questions such as whether classical logic is *the* one single 'true'/legitimate/correct/etc logic, or whether there are other logical systems that represent/describe in a better, perhaps more natural way, how humans think, and whether they should actually replace it or be considered mere variations or even extensions of classical logic. Throughout its history, possible extensions or variations of Aristotelian logic had been intuited – rather than considered –, namely regarding future contingents and modalities,[16] but until the beginning of the 20th century, with classical logic attaining its 'full form' with Frege and Russell, no one had seriously considered replacing it, namely for the reason that it did not comply with the way we actually think. Aristotle himself appears to have thought that bivalence is not universally applicable, with sentences about the future

the sets or classes, and shading of an area represents the presence of members (alternatively, shading can represent empty areas, with, say, an X representing non-empty areas).

[13] Note, for instance, that the set-theoretical operation of intersection 'is similar' to the logical operation of conjunction (the intersection of A and B, $A \cap B$, is simply the set of all the elements which belong to A and B, i.e., $A \wedge B$), and the set-theoretical operation of union 'is similar' to the logical operation of disjunction (the union of the sets A and B, $A \cup B$, just is the set of all the elements that belong to A or B, $A \vee B$). In fact, we can say that the laws of the algebra of sets have their 'similar' laws of the algebra of propositions (see Appendix, Fig. 12). The problem lies in what exactly 'is similar' means.

[14] To say that man is mortal is to predicate man (subject) of mortality, or the property or attribute of being mortal. Predication is connected to membership in an obvious way: for instance, to say that an object x has the property (also: attribute, predicate) P, i.e., Px, is the same as to say that x is a member of the set of objects that have the property P (formally, $\{x : Px\}$). The relationship between membership and predication/attribution of attributes becomes crucial in the case of sets the number of whose members is too large or even infinite, that is, in cases in which it is impractical or even impossible to actually list the members of a set. Note that the mathematical concept of infinity was precisely the motivation behind the constitution of set theory as a mathematical discipline (see Ferreirós, 2000).

[15] See for instance Russell's paradox, the class of the classes that are not members of themselves, which seems to contradict these two principles. Note that this, and other similar paradoxes (e.g., the liar paradox), was precisely the immediate motivation of changes and variations to this 'classical' set theory here being summarily described; the solutions found for the avoidance of these and related problems were type theory and category theory.

[16] Actually, it was Aristotle himself who first reflected on these issues.

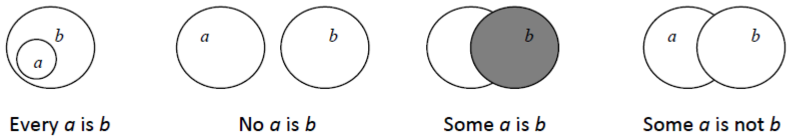

Fig. 3.1. Euler Diagrams 'Illustrating' the Four Aristotelian Forms of Statement.

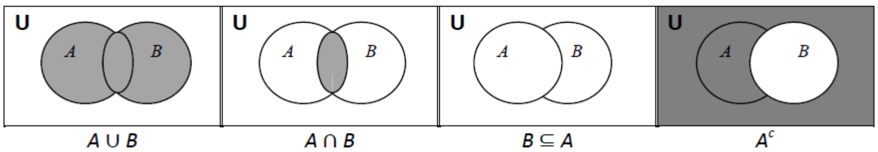

Fig. 3.2. Venn Diagrams 'Illustrating' Some Relations and Operations of Set Theory.

Fig. 3: Euler and Venn diagrams.

being undetermined, that is, neither true nor false,[17] and Hume (1748/1999) had fatally questioned the status of inductive logic, but logic was, well, *logical*, a science that had "come into a permanent state, which admits of no further alteration ... and indeed we do not require any new discoveries in Logic, since it contains merely the form of thought" (Kant, 1800/1885, pp. 10-11). The extraordinary developments undergone by logic until the end of the 19th century did not really change this Kantian depiction of it as the complete and finished description of the form of thought.

But if with Frege and Russell logic attained its full mathematical expression, this made it vulnerable to emerging attacks on the status of mathematics as a science: at the same time that the newborn set theory was struggling with serious paradoxes, threatening to leave mathematics without a rigorous foundation, the wholly a priori, ideal nature of mathematics (a stance known as platonism) was being questioned by constructivist approaches that saw mathematical objects as *constructible*[18] rather than merely *given*, and therefore requiring proofs that individual mathematical objects do indeed exist, i.e., that they are indeed constructible. In this perspective, it is not enough to assume that an object does not exist, thus entailing a contradiction: one has to actually prove that it does not exist, which is all but impossible if one

[17] This is the famous problem of future contingents firstly formulated by Aristotle in Book IX (19a30) of *On Interpretation* as follows: "A sea-fight must either take place to-morrow or not, but it is not necessary that it should take place to-morrow, neither is it necessary that it should not take place, yet it is necessary that it either should or should not take place to-morrow." (Transl. E. M. Edgehill) The important point here is that Aristotle seems to establish a new, intermediate truth-value that we can label 'indeterminate' or 'undetermined.'

[18] Namely as mental constructions.

is dealing with infinite sets. Take, for instance, the twin prime conjecture, stating that there exist infinitely many primes x such that $x+2$ is also a prime: because we have no method to decide whether this property is true or false for an arbitrary x contained in the interval from 0 to infinity, we cannot assert of an arbitrary x that it either does or does not have this property. This means that the law of excluded middle cannot be taken as an axiom, argued Brouwer (1907/1975, 1908/1975), the founder of this particular constructivist approach labeled intuitionism because it sees *intuition* as the (subjective) foundation of mathematics. In other words, the assumption that every mathematical problem has a solution is rejected, which means that parts of mathematics are unacceptable; given that for the intuitionists logic is secondary in relation to mathematics – namely, it is simply its description by the means of rules discovered a posteriori to be true of mathematical reasoning –, classical logic – or parts of it – is wrong.

Experience. That is how mathematics and its ally, logic, came to be plagued by assaults, the former seeing serious doubt thrown on its supposedly privileged, ideal status, and the latter seeing its a priori, normative status of the form of mathematical thought challenged. It took only a small step to generalize this problem firstly circumscribed to mathematics to *all* forms of thought: if mathematical thought is seen as subjective, empirical thought, and the logic attached to it is seen as subject to empirical laws, too, then it appears unacceptable to pretend to regulate in matters of human thought by appealing to the authority of (allegedly) a priori rules. Pandora's box was indeed wide open:

Why stick to the principle of bivalence, stating that each proposition is either true or false, when we know that many assertions are not necessarily one or the other – or are not provable as necessarily either true or false, as argued by intuitionism? There are many propositions (most of our assertions, probably) that are simply undeterminable regarding their truth value, at least in certain circumstances: so, why not introduce a third or even a fourth truth value? (See Figure 4 for an example of a truth-table for a 3-valued logic.) Why not infinite truth-values, as a matter of fact, given that an impressively large number of our assertions about the world suffer from radical imprecision or utter vagueness (e.g.: *X* is rich / old / beautiful / good / ...)?

Why should we stick to the principle of explosion, according to which from a contradiction follows anything, a property that is at first sight highly undesirable because making a logical system inconsistent and trivial, if there are non-trivial inconsistent theories? Many scientific theories seem to have this property (e.g., Bohr's theory of the atom), and there actually are true contradictions (e.g., the liar paradox) that make us see the principle of explosion as misguided.

Blindly relying on the material implication of classical logic is a source of bizarre statements that are nevertheless seen as valid. For instance, the formal rule $P \to (Q \vee \neg Q)$ originates an argument like "It is raining now in Mars. Therefore, either I am now in the cinema or I am now not in the cinema." as valid. What, if anything, has the premise got to do with the conclusion? In other words, what is their relevance regarding each other?

The cases above show that letting experience into logic implies taking *content* into account: every day we deal with information that is not only conflicting, but sometimes outright contradictory, and if this often forces us to review our beliefs, it sometimes also forces us to accommodate contradictions in our knowledge bases. This opening the door of logic to experience can as a matter of fact go as far as changing logic, instead of making changes in the body of a science, in order to allow

P	$\neg P$
1	0
u	u
0	1
Negation (NOT)	

P	Q	$P \wedge Q$	$P \vee Q$	$P \rightarrow Q$	$P \leftrightarrow Q$
1	1	1	1	1	1
1	u	u	1	u	u
1	0	0	1	0	0
u	1	u	1	1	u
u	u	u	u	u	u
u	0	0	u	u	u
0	1	0	1	1	0
0	u	0	u	1	u
0	0	0	0	1	1
		Conjunction (AND)	Disjunction (OR)	Conditional (IF, THEN)	Bi-Conditional (IF AND ONLY IF)

Fig. 4: Truth-tables for Kleene's three-valued logic ('*u*' is to be read as 'not known/indeterminable whether true or false' for undecidable sentences).

for empirical *observations* conflicting with classical logic: one example is quantum logics, devised to permit solutions for some problems in quantum mechanics, such as the Heisenberg uncertainty principle, according to which it is impossible to determine simultaneously both the momentum and the position of an electron, or any other particle, with any great degree of certainty or accuracy. With this end in view, quantum logics are characterized for the rejection of the distributive laws.

Thus, some of these 'revolutions' in classical logic were motivated by advances in empirical science, but mathematics itself was not immune to this 'empiricism'; for instance, the immediate motivation behind fuzzy logic, characterized by infinite degrees of truth (in the interval [0,1], where 0 is absolute falsity and 1 absolute truth), was the empirical fact that belongingness to a set (i.e., the membership function) comes in degrees. This, the idea that sets are 'fuzzy' rather than 'crisp,' is a major foundational intuition of fuzzy set theory (FST), and its empirical basis is obvious: for instance, dogs clearly belong to the set of animals (say, their degree of belongingness is 1), and rocks clearly do not (their degree of belongingness is 0), but what about bacteria? That is to say that 'being an animal' is not an all-or-none property, but that we can actually speak of degrees of being an animal; this makes 'being an animal' a *fuzzy* property (or predicate), and fuzzy *logic* was motivated precisely by the belief that 'true' and 'false' are also fuzzy properties. One can say for instance that a proposition such as "A bacteria is an animal" is *less true than*, or *not as true as* the proposition "A dog is an animal", which is itself *absolutely true*.

As said above, issues to do with modalities (necessarily; possibly) and temporal properties had been intuited since as early as Aristotle himself laid the foundations of logic, and the accommodation of these and other properties related to propositions all in some way having to do with experience (belief; knowledge; deontology; intensionality; etc.) were more or less gracefully added to classical logic as *extensions* that permitted to go beyond mere formal deduction, but the above cases stand for *alternatives* to it. Intuitionistic logic (characterized by the rejection of the law of excluded middle and also of double negation), paraconsistent logics (rejection of the principle of explosion), multi- and infinite-valued logics (rejection of bivalence), and relevance logics (rejection of the monotonicity of entailment) are, among other less well-known examples, some of the many logics that claim to be able to *replace* classical logic, altogether or in part (see Appendix, Fig. 13, for a summary view of the most important extensions of and alternatives to classical logic).[19]

3. Redefining Thought & Logic

With this scenario in front of us, what to make of the Wason Selection Task and the abundant research work it motivated (see below) around the humble material implication of classical logic? What to make of canonical schedules for the assessment of psychotic disorder heavily or entirely relying on 'crisp' sets? Is it that the psychology of thought lost touch with the rest of the sciences to which logic matters, or is it that these extensions, and especially the alternatives to classical logic are simply deemed unthinkable to be of any interest? If so, why? It indeed seems that there is something rotten in the kingdom of the psychology of thought.

At a time when non-classical logics and 'non-classical' set theory[20] are – slowly, one must admit – being applied to fields of the social sciences (e.g., the application of FST in studies of correlations – intersection – between poverty and depression; see Smithson & Verkuilen, 2006, for further examples and bibliography), and of psychology in particular (cf. Smithson & Oden, 1999, for abundant studies), one may legitimately wonder why the psychology of thought still sticks to the classical counterparts. This question is the more legitimate in that thought, as seen in the first part of this text, somehow comprises mental 'lower-order' processes that are being the subject of study in alliance with these non-classical developments: for instance, abundant work has been done on sensation and perception, at the very basis of thought as the source of its raw material, namely as far as the formation and categorization of representations is concerned; this work (see Smithson & Oden, 1999, p. 570) seems to support the idea that we categorize objects by representing them in semantic memory as fuzzy sets.

This is to say that the very initial steps of what we may see as higher-order mental processes, thought in general and reasoning in particular, are (or may be) more often than not characterized by inevitable vagueness and imprecision: we form vague and imprecise concepts that are pigeonholed in also vague and imprecise categories (e.g., tall, short, old, bald, beautiful; animal, bird, insect, fruit, vehicle, furniture; etc.), and it is from these that we then proceed to the other, higher or more complex thought processes, associating and combining. As the founder of fuzzy set theory put it, "such imprecisely defined 'classes' play an important role in human thinking, particularly in the domains of pattern recognition, communication of information, and abstraction" (Zadeh, 1965, p. 338), precisely those that at the same time were being recognized as major areas of study of the then emerging cognitive psychology (cf. Neisser, 1967). It is thus not surprising that fuzzy logic, the logic of fuzzy sets, found important applications in domains of (also) human information processing dealing with continuous variables, such as the linguistic variables (e.g., young/old age) that characterize the natural languages.

[19] The reader interested in more introductory material to the topic of non-classical logics is referred to Priest (2008) and Haack (1996). Although accessible, an understanding of the material in the former work requires a solid knowledge of at least classical logic (a short introduction to this is also provided), as it is a text of a rather technical nature; the latter makes in great measure for a philosophical reading. Both texts provide abundant bibliographies.

[20] Note that while fuzzy logic is claimed to be an *alternative* to classical logic, FST simply *extends* set theory, namely by generalizing the concept of set. Thus the inverted commas when referring to FST as 'non-classical' set theory. The same rationale is applied in the case of rough set theory (see below), as rough sets are *approximations* of the crisp sets of 'classical' set theory.

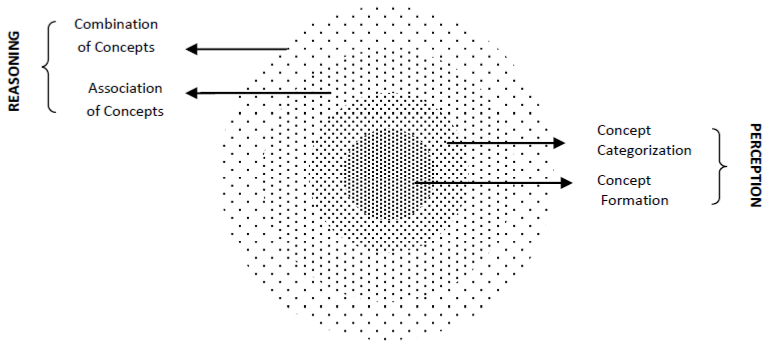

Fig. 5: A fuzzy hierarchy of thought processes involving concepts.

Should we then be expected to comply with the rules of classical logic and set theory, entirely based on crisp categorizations and bivalence? If not, how far can we stray away from them, and still be said to think logically, i.e., how far can alternative logics and variant set theories deviate from classical logic and set theory and still be considered logical? On the other hand, how much unaltered can classical logic remain if it is to be actually thinkable?

Let us then approach these issues by starting with the first steps of thought, and then moving on to the other, higher or more complex thought processes. In the spirit of FST, and for ease of treatment, I am dividing the four major thought processes very fuzzily into two major divisions, *perception* and *reasoning*, the first, in turn, comprising a fuzzy hierarchy from concept formation to categorization of representations, and the latter comprising a fuzzy hierarchy from association to combination of representations. Because concepts belong to the most studied representations, this hierarchy contemplates them, but it could just as well contemplate any of the 'mental stuff' introduced above. The diagram in Figure 5 might help visualize this fuzzy hierarchy.

3.1. Perception and Logic

By now, I think, it is very much obvious that the contemporary relations between psychology (in particular, psychology of thought) and logic cannot be seen as those of master and disciple, with one prescribing to the other, but of cooperation or even mutual influence. For instance, as reported by Smithson & Oden (1999), if fuzzy logic motivated a number of studies in psychology, these have, in turn, contributed to elucidate matters in FST and logic (see p. 570). This should not be surprising, as the very foundational notion of set theory, that of *set*, roots in perception and thought processes: for Cantor, the father of set theory, a set is "a collection into one whole of definite, distinct objects of our perception or our thought, which are called the elements of the set" (Cantor, 1895; translation from the German by Kneale & Kneale, 1984, p. 439). If, as seen in the words of Zadeh above, FST was already an extension of set theory motivated by a better understanding of human perception, rough set theory (RST), a more recent development,[21] makes this import of perception if anything even more salient:

> We can say that in Rough Set Theory, data are perceived by means of *anticipated hypotheses*. In other terms, the set A of observable properties that one decides to use, prepares a *perception grid* to be filled by the data deriving from our observations of $\mathbf{U}$ [universe]. This means that, since the very beginning, one cannot speak of mere data, but must speak of filtered data: data that one records by means of a well defined and specific perception grid. Data, as such, are Kantian 'noumena' that we perceive only within a categorization operation. (Pagliani & Chakraborty, 2008, p. xxxv)

And if this were not enough to allow us to realize how central perception is for RST, the authors actually go so far as to claim that RST is *a theory of perception* (their emphasis), in that it aims at interpreting observations by means of concepts, which, in turn, are "disjunctions of linguistically describable properties represented by unions of basic categories" (Pagliani & Chakraborty, 2008, p. lxii).

One of the motivations – perhaps the fundamental one – behind this recent development is precisely the realization that we more often than not have to do with incomplete information and with incompletely defined objects, and, as a theory of these two conceptions related to the incompleteness (or coarseness) of human information processing,[22] RST actually provides a polymorphic logic, at once intuitionistic, classical, modal, and three-valued (cf. *ibid.*, p. lxiii). The debate opposing logicism and psychologism has lost meaning: RST is *both* a logic *and* a (theory of) psychology, the description of the way we organize the universe from a cognitive point of view (ibid., p. lxv).[23]

We have gone full circle: we started by describing classical logic and set theory, saw some of their insufficiencies (more will be addressed soon) from the point of view of *actual* thought in that they are inseparably connected to experience, elaborated on how non-classical logics and set theory might help supplant those insufficiencies, to now conclude, with Pagliani and Chakraborty, that a conciliation of both classical and non-classical logic systems and set theory is actually necessary for an understanding, or representation of the way we do indeed perceive and think.

[21] RST introduces the idea of a *rough set*, constituted by two approximation sets, the set of objects that belong with certainty (i.e. necessarily) to the set (the *lower approximation*) and the set of those objects that possibly belong to the set (the *upper approximation*), and by a non-empty boundary between these two cases consisting of the objects we cannot decisively classify as belonging or not to the set due to limitations in our knowledge. The approximation sets may, in turn, be crisp or fuzzy. Note that this is itself a *rough* definition of rough set; see, for instance, Pagliani & Chakraborty (2008, pp. lxv-lvi) for a rigorous, yet accessible, treatment; alternatively, consult the immediate source of RST, Pawlak (1982). The rough definition above is enough for our purposes, allowing for a rough conception of rough sets and RST; our interest falls upon the fact that these sets are built from observable, empirical properties of objects and thus depend on our perceptive structure characterized by limitations.

[22] FST and RST are complementary in their modeling of uncertainty in information processing: whereas the former deals with vagueness, the latter occupies itself with coarseness; their combination allows for more accurate accounts of imperfect information, namely by combining both kinds of sets into fuzzy rough sets and/or rough fuzzy sets. Although both FST and RST find their applications mostly in artificial intelligence and related systems, their starting point is human perception, which processes only what is pertinent or meaningful; put in other words, we only perceive what our sensorimotor structure can process (in technical jargon: affordances). Interestingly, this shows how artificial information processing systems are determined at root by the human cognitive (in)capabilities.

[23] As a matter of fact, the same can be said of FST, though perhaps with less emphasis.

These recent developments in set theory and logic merely reflect, and attempt to model, what we have known for long: that the formation and categorization of concepts, whether natural or formal (i.e., learned in a formal environment or in day-to-day experience), is not a smooth, all-or-none business, with innumerable (probably most) concepts defying crisp sets and some causing havoc in theories of categorization, and this often due to different ways of perceiving the world. Ancient Greek philosophy started the discussion on the topic – namely with Plato's *Euthyphro* and especially Aristotle's *Categories*, the Scholastics were literally obsessed with it, both for metaphysical and logical reasons, and much contemporary philosophical work is still being done on it (see Margolis & Laurence, 1999). Psychology, in particular, and the cognitive sciences in general, are well aware of this problem, and it has been the focus of much theoretical and experimental work (see Medin & Rips, 2005, for a succinct review).

Nevertheless, we do not require more than lay everyday experience to be aware of the fact that we are highly tolerant of vagueness and coarseness in categories.[24] But if we are willing to accept that *furniture*, for instance, makes for a rough or fuzzy set (a regular sofa is unproblematically seen as a piece of furniture, but what about a set of curtains?, or a rug?, a standing ashtray?, etc.), that is, if we are willing to deal with vagueness or coarseness when referring to pieces of furniture, we are not (so) willing – poetical license apart – to accept *other* kinds of coarseness or vagueness in categorization, let alone the shifting of the boundaries of categories, fuzzy or rough as they may be, when we perceive them to be *transgressive*. Transgression, in turn, will be tolerated when caused by ignorance, especially when 'appearance' – thus involving (mis)perception – is concerned (e.g., you are likely to simply correct someone who says that a whale is a fish, given that whales actually have many fish-like properties: they live in the water, have fins, etc), but toleration will not be so prompt when transgression is seen as caused by *deviant* categorization, and this particularly so if deviance is seen as rooted in thinking clinically labeled as abnormal.

But what is it that makes categorization in, say, schizophrenia deviant even from the viewpoint of FST or RST, which appear to be so tolerant in their conceptions of fuzzy and rough sets and categorizations? A large number of studies abundantly document the fact that semantic categorizations are mildly to severely impaired in individuals with schizophrenia: they not only show longer reaction times when categorizing entities of the world, but also make a significant number of errors in tasks such as property attribution (e.g., "Rats have teeth": True or false?), picture naming, generating defining attributes, etc (see for instance Chen, Wilkins, & McKenna, 1994; McKay et al, 1996; Tamlyn et al., 1992).[25] Moreover, neologisms, often incomprehensible and whose creation, frequently attempting to categorize 'new' entities, indicates a distortion of reality, are not uncommon in more severe cases. These impairments or plain abnormalities in semantic categorization can be

[24] I am here restricting the discussion to semantic categories, and specifically of the verbal (lexical) kind.

[25] Data can be conflicting; for instance, in a series of studies, Elvevåg and colleagues have reported intact or only mildly impaired semantic categorization in patients with schizophrenia (e.g., Elvevåg et al., 2002; Elvevåg, Heit, Storms, & Goldberg, 2005), especially when the concepts tested are rather familiar (mammals, for instance; see Storms & Elvevåg, 2010). These results are indeed surprising and conflict with most data available from both first and third person observation since the labeling of the condition (e.g., Bleuler, 1911/1950; Kraepelin, 1896/1919; Schreber, 1903/2000). However, the results reported by Elvevåg and colleagues might be due to a number of factors, namely lower severity of the condition in the patients tested and/or the antipsychotic medication commonly taken by the patients participating in their studies.

said to be *illogical* because the boundaries 'allowed' by both FST and RST, hazy as they might already be, are completely transgressed, that is, entities are categorized in sets in which their degree of membership is 0, either because they simply do not belong to a particular set (see, for instance, Tamlyn et al., 1992) or because they simply have no referent in the universe of discourse capturing reality[26] (e.g., Schreber's *rays* and his *fleetingly-improvised men*; see Schreber, 1903/2000; see also Chaika, 1974).

In this view, *illogical thought* occurs when thought processes markedly deviate from the shifting allowed by the tolerance regarding the sort of semantic vagueness and uncertainty that 'non-classical' (i.e., tolerant) set theory can cope with. But this merely means that currently these thought processes are beyond the descriptive powers of logic, that is, the tools of logic fail to describe in a rule-like manner these particular thought processes. This, however, does not mean that possible further developments in logic will not be able to describe such thought processes as *logical*, perhaps much in the same way mathematics and physics were able to rigorously talk about chaotic systems and thus introduce chaos in formal thought.

This tells us that the expression *illogical thought* should not be seen as a value judgment; it simply refers to thought that is currently not describable by logic in its present state. Clearly, this poses the problem of how much 'tolerance' we can exercise and still be able to see deviant thought as (logical) thought, but this problem need not be different from the extent to which chaos can be mathematically determinable and still be seen as chaos. One thing is clear, at least, and that, I hope, is that bivalent, 'classical' set theory should not be seen as a descriptive – let alone normative – instrument of human thought processes, namely regarding the basic processes of formation and categorization of concepts; both FST and RST, more recent developments of set theory with firmer foundations on perception, allow for the limitations that characterize our basic thought processes while accounting for them in a rigorous way. We await possible further developments in set theory with great interest: to take place, they may well be a powerful utensil to understand the formation and categorization of concepts both in individuals with psychotic disorders and in professionals whose work tests the boundaries of these thought processes (artists in general, and poets in particular, come to mind), who, of course, are members of sets that can intersect.

3.2. Reasoning and Logic

Let us now turn our attention to the higher processes of thought, association and combination of representations. As seen, these are inseparable from the more basic thought processes, formation and categorization of representations, but for ease of treatment we shall consider the latter sufficiently discussed above, where we took set theory as their description in the logical perspective. Now, logic will be our main instrument, and we shall concentrate on the various operations conceived by logic as descriptive of human reasoning. Above, in the previous section, we saw how thought processes can be seen as illogical; that task shall be carried on in this section, but another, that of finding out what and when logic is unthinkable, shall also be undertaken. Because it is an intermediary thought process between the formation and categorization of representations in general and concepts in particular

[26] Of course, what this universe might be, that is a rather tricky metaphysical question, but, possible worlds apart, logic cannot indulge in such considerations, at least if it does not want to be seen as keeping the cake and eating it.

and their combination, commonly seen as reasoning proper, our attention shall fall firstly on association.

3.2.1. Association of Representations

If we take the five logical operators or connectives introduced above, to wit, negation (NOT; $\neg$), conjunction (AND; $\wedge$), disjunction (OR; $\vee$), the material conditional (IF ... THEN; $\rightarrow$), and the biconditional or material equivalence (IF AND ONLY IF; $\leftrightarrow$), we easily see that they can cover a vast proportion of our reasoning processes, i.e., the ways in which we manipulate representations, namely as regards their association and combination. This is especially so for propositions, but we can hypothesize that other kinds of representations (images, concepts, ideas, etc) also apply them; this is in particular true of concepts in that they can be seen as 'compressing' or 'condensing' propositions (e.g., the propositions "x is an animal", "x is a mammal", "x cohabits with humans", "x is a canine", "x barks", "x is often a pet", among many others, are 'compressed' or 'condensed' in the concept *dog*).

This can be at two levels; firstly, complex propositions are composed by two or more atomic propositions joined by logical connectives:

e.g.,

(1) "Brenda finished reading the book AND took it back to the library."

is a complex proposition composed of the two atomic propositions,

(2) "Brenda finished reading the book."

and

(3) "Brenda took the book back to the library."

Thus, if propositions (2) and (3) are symbolized as P and Q respectively, proposition (1) is then symbolized as $(P \wedge Q)$ if we want to take its internal structure into consideration; if we are interested in (1) as a single proposition without regard to its internal structure,[27] we can symbolize it as, say, R. On a second level, we can combine propositions, atomic or complex, with each other; for instance, we can combine proposition (1) with proposition (4)

(4) "Brenda was not involved in the bank robbery."

as proposition (5)

[27] In which case we should write "Brenda finished reading the book and took it back to the library", i.e., we write *and* as a grammatical connector instead of as a logical connective – in which case we write AND. Note, however, that this is not a universally followed convention, and the proposition "Brenda finished reading the book and took it back to the library" can, and often is, read as a complex proposition.

(5) "IF Brenda finished reading the book and took it back to the library, THEN she was not involved in the bank robbery."

If proposition (4) is symbolized as S, proposition (5) can, in turn, be symbolized as $R \to S$, or, if we choose to see (1) as a complex proposition, $(P \land Q) \to S$.

The same two levels can apply to concepts: at a first level, by taking concepts as expressing complex propositions, we can represent their internal structure by means of connectives. All five logical connectives above can be applied, but for clarity let us begin with conjunction; for instance, a concept like *dog* can express the conjunction of several propositions (see above), namely propositions expressing predicates or properties (e.g., the set of propositions $\{P_1 \land P_2 \land P_3 \land ... \land P_m\}$ can express the concept C_1). At a second level, we can associate and combine concepts among each other by means of the five logical connectives; for instance, we can symbolize the fact that concept C_1 is equivalent to concept C_2 as $C_1 \leftrightarrow C_2$.

Back to the logical connectives after this necessary excursion into logical symbolization for formal manipulation, classical logic will claim that the five operators above cover *all* our thought processes, constituting the operations of *the* syntax of thought, but the same motivations that lay behind the developments in 'non-classical' set theory also prompted extensions and even alternatives to classical logic; as seen above, neither a 'strict' bivalence is appropriate to describe the ways we think, nor do the five operators of classical logic capture all the linguistic modalities in which our thoughts can be expressed. For instance, we may think that a proposition P is necessary (representable as $\Box P$) or that some fact or property, say F, is necessarily true of an entity x (i.e., x necessarily has property F, representable as $\exists x(\Box Fx)$, which is the same as to say that it is not possible that x does not have property F, $\exists x(\neg\Diamond\neg Fx)$). Or we may want to express obligation, time, knowledge, etc. Furthermore, we may have to deal with incomplete information, true contradictions, and perhaps even with the puzzling empirical facts of quantum mechanics. All this – the introduction of modal operators; the acceptance of many or even infinite degrees of truth; the rejection of pillar principles of classical logic and the consequent absence of one or some of its theorems or valid inferences – introduced changes in both the syntax and semantics of classical logic, accounting for *new* logics.[28] That is to say that these new logics, whether extensions or rivals of classical logic, were deemed necessary to describe the various ways we think beyond classical deduction (see above; see also Fig. 13 in the Appendix).

As a matter of fact, we do not require all the five logical operators above, as most can be translated or expressed by the others; for instance, the material conditional, $P \to Q$, can be expressed as $\neg P \lor Q$. For classical logic, the following operators are the elementary ones: $\{\neg, \land\}$, $\{\neg, \lor\}$, or $\{\neg, \to\}$. These are thus complete sets by means of which classical logical calculus is said functionally complete (e.g., Wernick, 1942). On the other hand, many more operators beyond these five can be envisaged in a syntax, but again, they can all be expressed by the five commonest operators, which in turn can be expressed by only two or at most three, to wit, negation (necessarily) and conjunction and/or disjunction and/or the material conditional.

Interestingly, the set of logical operators or functors $\{\neg, \land, \lor\}$, commonly considered the complete set, describes what in this text is seen as the thought process

[28] Note that this means that, in the case of logics claiming to be alternatives to classical logic, what is involved is not merely a change of meaning (due, for instance, to mistranslation) of the logical connectives (see Haack, 1996).

of association: our concepts or representations are not commonly categorized in singletons, sets of exactly one single member, being associated with other concepts that share with them one or more properties; moreover, sets are in turn related to other sets as, for instance, subordinates of superordinate categories (e.g.: {Siamese cats} ⊂ {cats} ⊂ {felines} ⊂ {mammals}, etc). Negation, conjunction, and disjunction are enough for the description of these associations: some object O_i (that is, its concept) is associated with others belonging to the same set or class (e.g., let tabbies $= O_1$, Siamese cats $= O_2$, Manx cats $= O_3$; then they can be associated as: $\{O_1 \wedge O_2 \wedge O_3\}$ are cats), and some object that is not an O_i, i.e., $\neg O$, is excluded from that class, thus accounting for a disjunction of sets and memberships. In turn, cats, as concepts, are connected to a potentially large number of other concepts through more or less obvious properties (has fur; has claws; eats fish; is a finicky gourmet; is yellow/black/...; is clean; purrs; etc), thus making for an intricate, vast semantic network.

This, fluid as it may be as described by both FST and RST, seems to lie at the very basis of how our semantic memory is organized, with conjunctions and disjunctions standing for ways in which we *activate* or *inhibit* the – associations of – concepts that constitute our knowledge basis, the main instrument with which we approach the world.[29]

In studies of human semantic memory, the model of a *semantic network* by Collins and Loftus (1975), based on a theory originally designed to be implemented in computer simulations of memory search and later adapted to human memory and given a psychological elaboration, has been highly influential. This model represents the *relatedness* of concepts (*semantic similarity*) by aggregates of paths or connections and the *semantic distance* between two concepts (nodes, in the model; see Fig. 6) by the length of the shortest path that connects them. Modeling the theory of *spreading activation* of human semantic processing, the model assumes that memory is organized according to semantic similarity, and that when a particular concept is activated (for instance, by priming), those concepts in the network that are more closely related to it will also be activated. Accordingly, the model predicts lower reaction times to the question "Is a sparrow a bird?" as compared to the question "Is a penguin a bird?", because the concepts *bird* and *sparrow* are more closely reciprocally related than the concepts *bird* and *penguin*, as an important property of the concept *bird* is *flies* and, whereas sparrows do fly, penguins do not. More important for our ends, the model predicts activation of only related concepts, i.e. nodes connected by a line, no matter how long the line might actually be and how many nodes might intermediate between the concepts; the activation of non-related concepts is – or should be – inhibited. This explains why we keep to the main topic of a conversation and make only somehow predictable changes regarding it. For instance, one may start talking about cats, and then move on to food via considerations on what finicky eaters cats are, and then move on to the cost of living due to the appalling prices of food, then to unemployment, money, interest rates,

[29] This is what explains the separation in this text between simple categorization, analyzed above, which can be described from the set theoretical viewpoint as a function of degree of membership, and association, which involves the logical operations of negation, conjunction, and disjunction. But neither the separation between categorization and association, nor that between logic and set theory is to be taken rigidly; as seen above, we could as well approach association with the operations of set theory corresponding to those of logic (see Appendix, Fig. 12). The choice here of logic instead of set theory is explained by the fact that this will allow for a smoother transition between the treatment of association and that of combination of representations, which requires a more strictly logical approach.

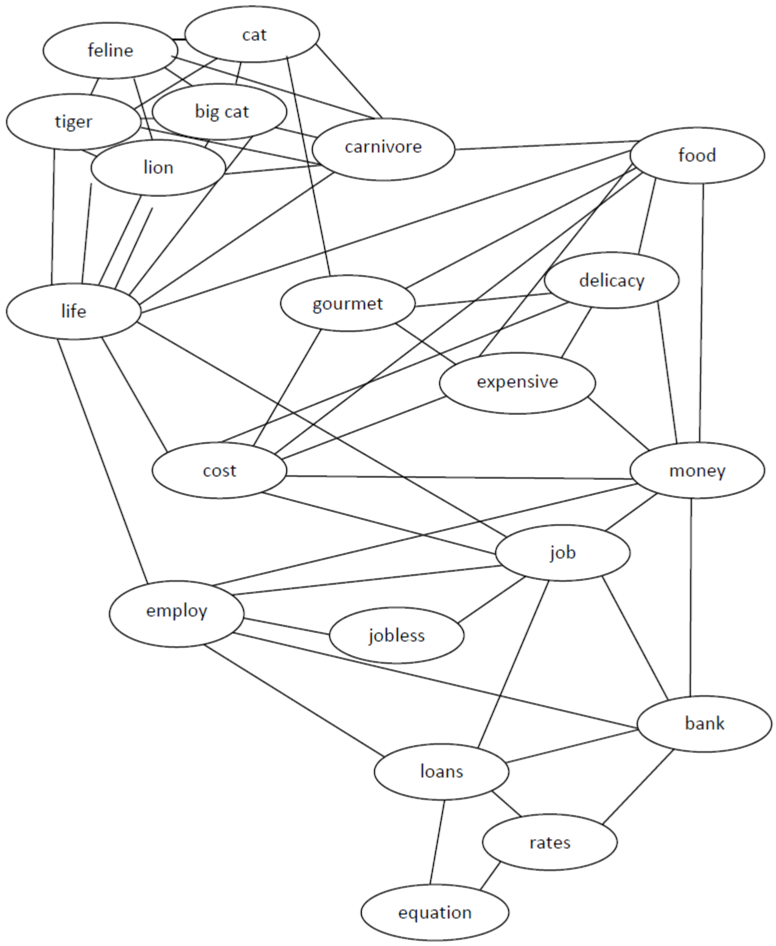

Fig. 6: How do you go from cats to equations? An adaptation of Collins and Loftus' model of semantic memory (cf. Collins & Loftus, 1975, p. 412).

and possibly end up discussing differential equations; to move suddenly, perhaps in the same sentence or clause, from cats to differential equations would be unexpected, because the semantic distance between the two concepts can be quite large and, for those not versed in differential equations (or in cats!), simply inexistent for lack of one of the concepts (see Figure 6).

In fact, Collins and Loftus' theory of semantic (and/or lexical) spreading activation is an AI-inspired refinement of one of the theories that helped launch linguistics as a science: Saussure, the father of scientific linguistics, already spoke of the *associative relations* that constitute a hidden network[30] of semantically or lexically associated concepts that are unconsciously evoked when one thinks of a particular concept (see Saussure, 1916/1977). This, in turn, is the linguistic expression of an idea that had long before been explicitly stated in philosophical terms (Locke, 1690/1959; Hume, 1748/1999) and had found its way into experimental psychology

[30] Saussure spoke of a system.

(Ebbinghaus, 1885/1913; Jung, 1904-10/1973). This blitz-short history of the antecedents of Collins and Loftus' model are of interest for us because it was precisely within this psycho-linguistic associationism that the clinical concept of schizophrenia first became more distinct as a *loosening of associations* (Bleuler, 1911/1950), an idea that was first conceived and experimentally put to the test by Jung, who saw Kraepelin's (1896/1919) concept *dementia praecox* as rooting in chronic pathological complexes of associations (Jung, 1907/1960). Many contemporary studies approach the clinical condition of schizophrenia from this Bleulerian perspective supported by a spreading activation model; this approach constitutes what we can call the dyssemantic hypothesis of thought disorder[31] (DHTD; see McKenna & Oh, 2005, chapter 7).

There are actually many studies and individual theories that can be seen as a part of DHTD, the common point among them being the hypothesis of a failure in the organization and/or activation of the associations that compose an individual's semantic memory. For instance, Maher (1983) proposes that the spreading activation may be increased (or inhibition decreased) in patients with schizophrenia, whereas other studies (e.g., Goldberg & Weinberger, 2000; Leeson et al., 2005; Niznikiewicz et al., 2010; Waters et al., 2003) emphasize inhibition or the very organization of the patients' semantic memory as the cause of the abnormal associative processes behind the bizarre or plainly aberrant speech productions that are often encountered in patients with schizophrenia, and that have been considered a symptom of the condition from the very beginning of its description: for instance, Kraepelin already saw expressions such as "I have a suspended appetite", "I have voluntary disease of the eyes" and "they are threaded at the head" as a gliding off into side-ideas contributing to what he called akataphasia, or "derailments" in the expression of thought in speech (Kraepelin, 1896/1919, pp. 70-1).

Examples of impaired semantic associations in formal thought disorder abound in the literature on schizophrenia, but the study of Tamlyn and colleagues (Tamlyn et al., 1992) is particularly interesting, because it uses tasks to test semantic association that were devised by Collins and Quillian (1969) in the then developing paradigm of spreading activation of semantic memory. Tamlyn and colleagues administered a variant of the task devised by Collins and Quillian (1969), in which subjects have to answer 'true' or 'false' to sentences like "An elm is a plant." The study by Tamlyn and colleagues was already cited above in the section on categorization: the subjects, patients with chronic schizophrenia, showed abnormal categorization by answering 'true' to sentences like "Owls have blades" and "Screwdrivers have a profession." This already shows evidence of bizarre associations, but the ways in which the patients justified their answers were better revealing of these anomalous thought processes; for instance, answering 'true' to the sentence "Crows are in charge of ships", a patient justified this with the statement "In the crow's nest."

These associations are indeed bizarre, to say the least, but are they, from the viewpoint of logic, illogical? That is, are we in any way entitled to see as illogical what in psychological and psychiatric terms is said to be aberrant, abnormal associations symptomatic of some awry, "private logic" that seems to work differently from Aristotelian logic (e.g., Akiskal, 2008)? We saw that this level of thought is expressed, in logical terms, by means of the operations of negation, conjunction, and disjunction; therefore, in order to say that these thought processes in schizophrenia

[31] More specifically, of formal thought disorder.

are illogical we must first see in which way they go against the truth-tables for these connectives.

From the viewpoint of classical logic, they can indeed be seen as illogical thought processes, in that they do not comply with the complete sets $\{\neg, \wedge\}$ and $\neg, \vee\}$, i.e. they activate (do not inhibit) unrelated concepts and inhibit (do not activate) related concepts. We begin with the logical operation of conjunction. Let us take related concepts as having the same 'truth' value 1, and unrelated concepts as having either different 'truth' values ($\{1, 0\}; \{0, 1\}$) or as having the same value 0; furthermore, let us represent activation (= disinhibition) by 1 (the concepts are related, so the association should be activated/disinhibited) and inhibition (= non-activation) by 0 (the concepts are not related, so association should be inhibited/not activated). By activating unrelated concepts (e.g., $\{0, 1\}; \{0, 0\}; \{0, 1, 0\}$), and by inhibiting related ones ($\{1, 1\}; \{1, 1, 1, 1\}$; etc), we can say that patients with schizophrenia, from the viewpoint of logic, are illogical, namely in that they operate in a way completely contrary to a computer implementation of the model of spreading activation of a semantic network (e.g., the Teachable Language Comprehender of Quillian, 1969). We can easily see this if we revisit the truth-table for conjunction in Figure 2 above and compare it with what we can call '*schizophrenic conjunction*' (see Fig. 7).

C_1	C_2	$C_1 \wedge C_2$
1	1	0
1	0	1
0	1	1
0	0	1

Fig. 7: Truth-table for 'schizophrenic conjunction.'

We now move on to the logical operation of disjunction. Let, again, related concepts be represented by 1, and unrelated ones by 0 alone or combinations of 1 and 0. Let us suppose that, in a lexical decision task regarding semantic relatedness, the subject has to choose whether to activate both concepts presented (i.e., $\{1, 1\}$), one of the concepts alone ($\{1, 0\}, \{0, 1\}$) or none (i.e., $\{0, 0\}$), that is, s/he has to inhibit any activation. Then, '*schizophrenic disjunction*' can be described as follows (see Fig. 8): the patient of schizophrenia fails to choice-activate concepts that are related and fails to choice-inhibit those that are unrelated. By comparing Figure 8 with Figure 2 above, we see that we are again facing the opposite of what the truth-table for the classical operation of disjunction prescribes.

C_1	C_2	$C_1 \vee C_2$
1	1	0
1	0	0
0	1	0
0	0	1

Fig. 8: Truth-table for 'schizophrenic disjunction.'

After this, the ways in which patients with schizophrenia may be seen as thinking against logic when negation is concerned are obvious: they may simply fail to distinguish a concept from its contrary (e.g., *alive* – *dead* [= *not alive*]; *existent* – *inexistent*; *natural* – *artificial* [= *not natural*]), that is, they fail either to activate the disjunction of two antonymous concepts, or they do not inhibit their association as synonymous; in either case, classical logical negation fails. They thus fail to comply with the pillars of classical logic, the principles of contradiction (PC) and of excluded middle (PEM). For instance, when asked in a study how s/he did like it in the hospital, a patient with schizophrenia replied to the interviewers as follows (see McKenna & Oh, 2005, p. 10):

> Well, er… not quite the same as, er … don't know quite how to say it. It isn't the same, being in the hospital as, er… working. Er… the job isn't quite the same, er… very much the same but, of course, it isn't exactly the same.

While in his/her answer the patient applies negation correctly from the grammatical point of view ("[I] don't know quite how to say it."; "It isn't the same…"; "The job isn't quite the same."), s/he fails to clearly distinguish an affirmation and its negation, saying "[the job is] very much the same but, of course, it isn't exactly the same." The failure to comply with the PC, and with its complement, the PEM, is obvious. Schreber (1903/2000, p. 26) provides a more refined example: while recognizing the distinction between a concept and its opposite or antonym, he spoke of the antonymy of some psychologically charged concepts (*reward* = *punishment*; *poison* = *food*; *juice* = *venom*; *unholy* = *holy*) as a euphemism. In other words, what Schreber called the "basic language," used by the voices in his auditory hallucinations, was characterized, in logical terms, by the non-distinction between a concept and its negation or opposite; for instance, if to the concept *reward* the 'truth' value 1 were to be attributed, this same value would be attributed to *punishment*, and if to *unholy* the 'truth' value 0 were attributed, then *holy* would also be valued as 0. The table for '*schizophrenic negation*' (Fig. 9) illustrates this disregard of bivalence.

C	$\neg C$
1	1
0	0

Fig. 9: Truth-table for 'schizophrenic negation.'

Without the PC and the PEM, we are undoubtedly confronted with illogical thought; we can thus say, from the viewpoint of classical logic, that the thought processes of association in patients with schizophrenia are indeed illogical. But we saw above that an immediate motivation of extensions and alternative logics, namely of logics with three or more truth values, is precisely the fact that we often (or more often than not) have to deal with vague, incomplete, or even contradictory information. Nothing could be truer of semantic information, contrary to what the model of Collins and Loftus (1975) might suggest. Take, say, the adjective *bald*: when is one exactly bald? One who has no hair at all due to alopecia totalis is said to be bald, but one who has only a bald patch on the top of the head, a common sight in androgenic alopecia, is also referred to as bald; diffuse hair fall can also cause a person to be called bald, though there is hair in the scalp; and if one chooses to shave one's head, one is also called bald, though one might have a scalp completely sewn with active hair follicles. Does this entail that one can be bald and at the same time not bald? And does only hair in the scalp count for baldness, or can we say of a person that s/he is bald when s/he has no body hair? What about other mammals, that is, fur-covered animals: is, say, a Sphynx, a cat also known as Canadian Hairless because it has no coat (hair) whatsoever, also bald? It is hairless, and is not hairless a synonym for bald? So, can we or can we not associate baldness with hairless cats? Featherless birds? Leafless trees?

Examples like these – and some far more problematic – abound in any semantic network associated with a natural language, and we should thus perhaps opt for a multi-valued logic when assessing the logicality of our semantic associations. We could, for instance, opt for a three-valued logic in which the presence of an undetermined or vague concept (with truth value *i*) hinders either activation or inhibition of semantic associations. In fact, Smiley (1960) has one such logic ready at hand for us, where *i* is read 'undefined' or 'truth-valueless" (see Fig. 10).

C	$\neg C$
1	0
i	i
0	1

C_1	C_2	$C_1 \wedge C_2$	$C_1 \vee C_2$
1	1	1	1
1	i	i	i
1	0	0	1
i	1	i	i
i	i	i	i
i	0	i	i
0	1	0	1
0	i	i	i
0	0	0	0

Fig. 10: Negation, conjunction, and disjunction in Smiley's three-valued logic.

It is true that in the clear-cut cases ($\{1\}$ and $\{0\}$ for affirmation and negation; $\{1,1\}$, $\{1,0\}$ and $\{0,1\}$, and $\{0,0\}$ for conjunction and disjunction) the tables function as in old classical logic, but then, when does one have clear-cut concepts? Not often, that is certain. Adopting these truth-tables would entail as a consequence that the assessment of what psychology and psychiatry often hastily see as aberrant, illogical associations, are in fact not illogical, or are only so from the restricted viewpoint of classical logic. Clearly, adopting a logic like Smiley's in the constructions of programs simulating the organization and activation of human semantic memory might be simply counterproductive, but this is what gives us the cue to our next and final discussion, viz., when is logic unthinkable.

3.2.2. Combination of Representations

Let us then now concentrate on the combination of representations, the thought process commonly known as reasoning, but which in this text we take in a less restricted sense (see above). We are now in the terrain where logic is applied in full: the combination of representations, besides applying the logical operations of negation, conjunction, and disjunction, also operate with the material conditional and with material equivalence, that is, the combination of representations uses the set of connectives $\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow\}$ claimed by logic to be capable of expressing our thought processes. In this terrain, logic has thus a lot to offer (and to say!) for the understanding of thinking, human or 'artificial,' namely in what respects deductive reasoning, or reasoning involving inference. After all, logic aims precisely at establishing the forms or laws of valid inferences, seeing itself as the science of the arguments or argument forms that we are supposed to adopt – or should adopt – when thinking, and though there are those who entirely dismiss the claim that logic describes human reasoning, it is more commonly believed that, at least to some degree, it does so. But does it? And if so, to what extent? And if not, why not?

The contemporary, perhaps most influential discussion around reasoning, or deductive thought, arose in great measure from results reported by Peter Wason concerning performance in a selection task involving rules of the kind "If P then Q", that is, conditional rules (Wason 1966, 1968). Cutting a long story short (but see, for example, Stenning & van Lambalgen, 2008, for details), Wason verified that only a mere 4% of subjects tested in the Wason Selection Task conformed to classical inferential reasoning, and concluded that most subjects appeared not to have attained what Piaget called the stage of formal operational thought (Wason, 1968) – in other words, they were illogical or irrational.

But were they? A whole regiment of researchers did/does not think so, with harsh criticism falling on the versions of the tasks conceived by Wason, namely on the irrelevance or inadequacy of the content and/or context involved (again, see Stenning & van Lambalgen, 2008). Notwithstanding, nearly all the critics of the Wason's task and conclusions completely ignore the developments in logic we approached above; commonly, they either stick to old classical logic in their approaches to the psychology of reasoning, or simply dismiss logic as altogether irrelevant for it. Neither group is correct, the former for the reasons seen above regarding the development of logic (and for further reasons to be discussed in what follows), and the latter for not acknowledging that logic can contribute to an understanding of thought processes, not the least because it claims to be their description, if not their norm.

Let us then turn to classical logic and the claim that it describes *actual* human reasoning. It is not uncommon to see logicians, mathematicians, and philosophers at large claiming that classical logic expresses the structure of our thought, thus following in the footsteps of pre-cognitive psychology logicism represented by names such as Frege (1884/1953), Husserl (1900/1970), or Hilbert (1931/1996). Boole (1854/1958) represents an interesting position in that he defends that while logic is the study of the laws of thought, it treats only of the laws of *right* reasoning, which "form but a *part* of the system of laws by which the actual processes of reasoning, whether right or wrong, are governed" (Boole, 1854/1958, p. 409). The interesting aspect lies more precisely in the fact that Boole not only saw that while being laws, the *laws of right reasoning* are not like physical laws (they can be, and often are, violated, without, however, being affected in their necessary character), but also that there must be *laws of wrong reasoning* that explain both the violations of the laws of right reasoning and other erroneous reasoning processes. In fact, it has been known for long that at least some errors in reasoning, fallacies, follow rules of their own, besides deviating from or violating the laws that govern right reasoning; the medieval logicians, in particular, carried out a rigorous study of fallacious reasoning processes. But Boole, interestingly, did not see the study of error in reasoning as part of the scientific enterprise (cf. Boole, 1854/1958, p. 408), seeming to attribute error to some mysterious or in some other way unaccountable property of the human mind.

While Boole's (1854/1958) aims at being a science of the mind, "designed, in the first place, to investigate the fundamental laws of those operations of the mind by which reasoning is performed" (p. 3), and although he defends that such a science relies, too, on observation of the actual way people do reason, he distinguishes *this* observation-based science of the mind from *the* science of the mind: in the former, one may doubt whether a law such as Aristotle's *dictum de omni et nullo*[32] is or expresses a law of reasoning or not, but it undoubtedly expresses a law of the latter, that is, of logic. Because it is evident, one single instance of it suffices to accept it as a law of thought, or to assent to its necessary character as a law of the human intellect; moreover, this, and the other *logical* laws, can be expressed by means of the symbolic language of a calculus. Boole thus has the merit of being the father of symbolic logic, bringing this science to its essentially contemporary form. What Boole does not do is give a name to – and dismiss? – the *other*, observation-based science of the mind, which can observe whether logical laws are laws of actual reasoning, that is, whether people actually apply them *necessarily* when reasoning. This science is, of course, psychology. This omission is at least intriguing, as by the time Boole wrote his two major texts on logic (Boole, 1847/1948; 1854/1958), psychology had already shown its propensity for empirical observation and experimentation (e.g., Mill, 1829), and Mill (1843), a text that Boole seems to have known well, though avoiding the terms *psychological* and *psychology* – or, as he also called it, science of human nature or mental philosophy – in the Introduction, ends up making an abundant use of them, particularly in Book VI, where he all but makes logic depend on what today we call psychophysiology:

> the laws of mind may be derivative laws resulting from laws of animal life, and ... their truth, therefore, may ultimately depend on physical conditions; and the influence of physiological states or physiological changes in altering or counteracting the mental successions, is one of the most important departments of psychological study. (Mill, 1843, p. 591)[33]

The case of Boole – and, to a lesser extent (?), of Mill – is symptomatic of the position that, seeing the laws of logic as the laws of thought, and therefore in some way laws of psychology – at least to the extent that this is the science of the actual application of these laws –, still regards logic as untouchable, no matter how *unthinkable* its laws might actually be. The conceiver of the Wason Selection Task radicalizes this perspective. Wason's (1968) article sparked an astonishingly abundant research,[34] but little or no work was done on the thinkability of the rules and laws of logic, that is, the question "can we actually think this?" was not asked explicitly (enough). For instance, studies were done with illiterate subjects from cultural backgrounds profoundly different from the called western one (e.g., Luria, 1976; Scribner, 1997), with somehow expected results (for instance, subjects simply refused to reason without real acquaintance with the contents of the premises), but

[32] Fundamental principle of classical logic according to which what is affirmed or denied of a kind *M* (e.g., mammal) may be affirmed or denied of a subkind of *M* (e.g., cat).

[33] Page number relative to the 8th edition, published in New York by Harper & Brothers, in 1882.

[34] As a measure of the popularity of Wason (1968), let the number of results given by Google to the title of this article, "Reasoning about a Rule", suffice: around 177,000!

our interest should primarily fall on the claim that the *rules* of logic, commonly symbolically expressed or only with very little flesh to them, are actually rules of thought in the sense that they are actually thinkable.

Interestingly enough, the first reactions against the 'unthinkable' character of prevalent logical systems came, not from psychology, but from the field of logic, in particular of mathematical logic. In 1926, in his seminars, the Polish logician Łukasiewicz's criticized the existing axiomatic systems (i.e., mainly Frege, Hilbert, and Russell's logical systems) as not reflecting the way mathematicians make proofs; according to him, they do not follow axioms, rather using more 'informal' methods, such as following arbitrary assumptions in order to see where they lead. Motivated by this critique, Jaśkowski (1934) came up with an axiom-free system that was nevertheless provable to be equivalent to the established axiomatic systems. At the same time, and independently, Gentzen (1934-5/1969), also dissatisfied with the axiomatically-based logical systems at hand, according to him extremely remote from the deductive reasoning that is actually carried out in mathematical proof, developed what for him was a formal system that came as close as possible to actual reasoning; he called it "calculus of natural deduction." This virtually axiom-free formalism, or calculus, is known as *natural deduction* (henceforth: ND), as it is believed to reflect *natural*, that is, informal and intuitive, reasoning (cf. Prawitz, 1965/2006, p. 1), and, in place of the abundant axioms of the Hilbert-style systems, its basis is constituted by inference rules. These are divided into rules of introduction ($\wedge I, \vee I, \neg I, \rightarrow I, \exists I, \forall I$) and of elimination ($\wedge E, \vee E, \neg E, \rightarrow E, \exists E, \forall E$) for the connectives and the quantifiers (see Appendix, Fig. 14)[35] believed to reflect the way we think; for instance, inference rule (1.1) (see Fig. 14) reflects the fact that if you assume that A is true and that B is true, you can deduce that $(A \wedge B)$ is also true by introducing the connective for conjunction ($\wedge I$); assuming, in turn, that $(A \wedge B)$ is true, you may deduce by inference rule (1.2) that A (or B) is true by eliminating the conjunction ($\wedge E$).

Intimidating as it may appear to the uninitiated in logical matters (at least in a first encounter), this symbolic formalism is necessary, as ND is concerned with form alone when making trees (i.e., deductions from assumptions), which can be quite long and thus even more intimidating (see, for instance, Prawitz, 1965/2006, for abundant examples). This is what today is commonly taught as logic in introductory courses of the subject, and our main interest is to see whether these inference rules do indeed reflect or illustrate the ways in which we actually think when reasoning.

The first, often neglected, aspect, is that though we think with representations, of which abstractions in general are a part, we do not naturally (i.e., non-strategically) think with formal artifacts such as A and B standing for propositions or formulas; in a way that is still to be understood, our thought is connected to verbal discourse, not the least because we can verbally report (translate?) our thoughts, namely when communicating them. Secondly, language is, in a way that is also to be fully understood, intimately connected to culture, and we thus have the triangle thought – language – culture to deal with when addressing the issue of thought processes. Nevertheless, at the center of this triangle we can posit *schemas* that, though they operate with a natural language and the culture attached to it, are to some extent independent from them in that every production, mental and/or verbal, in a culturally imbibed natural language cannot fail to follow them if is to be counted as rational; in other words, these would be *universal thought schemas.*

[35] There are variations of these figures; here, I follow closely Prawitz (1965/2006), p. 20.

These are what Braine calls *primary skills* and it is through them, and through them alone, that according to him we can test people's reasoning skills: fail in a secondary skill, and all is fine, but fail in a primary skill, and you are irrational, seems to be the essence of Braine's proposal (Braine, 1990, p. 138). For our interests, what Braine calls *primary skills* corresponds basically to the rules of inference of ND, and constitute the inference schemas of the theory known as *natural logic* (henceforth: NL; see Braine, 1990, for the essentials of the theory).[36] NL began by contemplating only the propositional part of ND (i.e., propositional logic), and the work of Braine (1998) extends NL by introducing inference schemas for predicate logic. We shall call this extension *mental logic* (ML), though this label may contemplate only propositional logic; as a matter of fact, this label has been around for quite a while, too, and NL and ML seem to be interchangeable labels, with ML becoming the more orthodox one.[37]

Below (Fig. 11) are given the inference schemas of ML corresponding to the inference rules of natural deduction (see Appendix, Fig. 14), with the bonus of examples with natural propositions, as Braine calls them (Braine, 1990, pp. 140-1);[38] as the correspondence between them is not perfect, approximations are marked with *. Although Braine (1998) extends ML by introducing schemas for predicate logic, the truth is that he does not consider the very basic rules of introduction and elimination for the quantifiers *per se* (see Appendix, Fig. 14, Rules (5.1), (6.1), (5.2) and (6.2) respectively). This is not Braine's fault; the fact is that these rules and schemas seem to require the other rules for propositional logic (e.g., the rule of inference for the universal elimination seems to require *modus ponens*, i.e., schema (3.2) in Fig. 11), and thus actually require implicit premises. In the list below, these implicit premises are marked with ⫳, and the inference schemas for the quantifiers are marked with ** to signal the fact that they are not considered in Braine (1998). Nevertheless, the reader should note that this already shows how both the inference rules of ND and the inference schemas of NL/ML are inadequate to represent basic thought processes.[39]

[36] Braine sees the *secondary skills* as heavily dependent on literacy and "language qua language", as he puts it (cf. Braine, 1990, p. 154). This is highly problematic, because, as Braine himself sees it (cf. Braine, 1990, p. 149), the use one makes of the logical connectives involved in what he calls the *primary skills*, available early to children and present even in pre-literate societies, is, in large measure or perhaps even entirely, inseparable from their meanings in a natural language, *too*. The distinction is thus rather opaque. Moreover, if what he calls primary skills are to *any* extent connected to a natural language, his inference schemas cannot be seen a priori as universal, an exhaustive empirical study aiming at verifying if the logical connectives for negation, conjunction, disjunction, and material implication are to be found in *all* natural languages, present as well as past, being necessary. On the other hand, if one sees the logical connectives as natural language connectives, imperfect as the equivalence may be, then logic is simply a formalization of natural language, and there is no more to it. As we shall see, quantificational inference will make this problem more acute.

[37] This in spite of there being another, supposedly completely different theory developed by Rips (1994) also called ML.

[38] The model also includes an associated reasoning program (see Braine, 1990, p. 142), but here we can skip it.

[39] The problem is more complex, however, as the propositional inference rules (1.1) to (4.2) seem to involve already quantification: when a child points to a dog and says "Doggy!", s/he is making an existential statement of the kind "There is a dog" (more explicitly, "There is at least one x such that x is an A."). This seems to entail that existential quantification precedes the other, propositional, inference schemas, but when we get to the schemas for quantificational inference, the propositional inference schemas seem to precede. I find this problematic. Burge (2010) does not (see p. 547): for him, though quantificational inference is in a way an elementary skill, propositional inferences are more basic and precede them.

But let us have a close look at these schemas said to be basic of human reasoning processes (see Fig. 11).

Much has been said about the gap between quantification in natural languages and in first-order logic (see Molczanow, 2004, for an overview), but I am here going to concentrate on the thinkability of the inference schemas, in particular of the quantificational inference schemas. By doing this, that is, by aiming at showing that these schemas do not reflect natural, i.e., basic and universal processes of thinking about quantification, I have no intention of dismissing the whole of modern symbolic logic, as, for instance, Molczanow (2002) does; I will be merely showing that some parts of it are in the borderline of the intersection between logic and thought, nearly – or actually – making for unthinkable logic (see Fig. 1).

Above, immediately before the list of the inferentional schemas of mental logic (Fig. 11) was introduced, I called the reader's attention to the fact that the rules for the introduction and elimination of the universal and existential quantifiers need, implicitly or explicitly, the rules for the connectives, which shows that they are not as primary as claimed. As we shall see, these rules are actually so complex that they simply do not reflect what we can call *basic* thought processes, as inference rules in ND are supposed to do. Instead of schemas reflecting natural basic ways of thinking, they are rules for artificial thinking processes deemed correct for inferences involving quantification; hence, they are artifacts. This, however, is not where their sin lies, given that artifacts can be thinkable, even at what for Braine is a primary level; their sin lies at their being simply unthinkable at a primary, basic level. But this is only one sin of at least two: their second sin lies in the fact that, contrary to the claims of ML, they are not *universal*.

We start with the second sin. Let us then start with the inference schemas of ND/ML featuring quantification. They are usually introduced in two pairs of distinct degree of difficulty or complexity: universal elimination (schema 5.2 above) and existential introduction (schema 6.1.), considered less problematic than universal introduction and existential elimination (schemas 5.1 and 6.2 respectively). We follow this order.

Universal elimination, also known as *universal instantiation*, because it involves the inference from a proposition stating that all things are / everything is so and so to an instance stating that a particular thing / something is so and so, describes the rather trivial inference that if everything is P or has the property P, then some particular thing is P or has the property P. We make this less trivial by defining a domain of discourse, x, and a constant c that can replace x. Formally, $\forall x Px \Rightarrow Pc$,[40] where $\Rightarrow$ symbolizes logical consequence and is read as "entails" or "implies."[41] In the natural proposition illustrating the schema above (Fig. 11) for universal elimination, the domain of discourse, our x, is cats, our constant, c, is

[40] The rather off-putting A^x_t in inference rule (5.2) in Fig. 14 simply represents the substitution of all the occurrences of the term x by the term t with respect to a formula A. Let P above stand for a formula; then we can represent the substitution of x by a constant c as P^x_c.

[41] I am here assuming that *entailment* (represented by the symbols $\Rightarrow$, $\therefore$, or $\models$) is extensionally equivalent to *deduction* (represented by the symbol $\vdash$), but this is not exactly so, with, to begin with, entailment standing for a semantic consequence relation and deduction for a syntactic one. The distinction is important for the definition of the completeness and soundness of a formal system (a formal system is *complete* if for every set of formulas Φ and every well-formed formula φ of that system, if $\Phi \models \varphi$, then $\Phi \vdash \varphi$, and it is *sound* if for every set of formulas Φ and every well-formed formula φ of that system, if $\Phi \vdash \varphi$, then $\Phi \models \varphi$), but for the ends of this paper it is not essential, and both entailment and deduction are treated as logical consequence, unless otherwise stated. For an introductory elaboration on and refinement of the concept of logical consequence, see Blanchette (2001).

(1.1)	*P*; *Q* *P* AND *Q*	There is a cat; There is an apple /∴ There is a cat and an apple.
(1.2)	*P* AND *Q* *P*	There is a chicken and a horse /∴ There is a chicken.
(2.1)*	*P*; *Q* *P* OR *Q*	There is a cat; There is an apple /∴ There is a cat or an apple.
(2.2)*	IF *P* OR *Q*, THEN *R* *P* OR *Q* *R*	If there is a duck or a goose, then there is a cherry; There is a duck or a goose /∴ There is a cherry.
(3.1)	Suppose *P* .... *Q* IF *P* THEN *Q*	There is a fox. If there is a fox then there is a wolf. If there is a wolf then there is a lizard. There is a lizard /∴ If there is a fox then there is a lizard.
(3.2)	IF *P* THEN *Q*; *P* *Q*	If there is a fox, then there is a nut; There is a fox /∴ There is a nut.
(4.1)**	*P*; INCOMPATIBLE NOT *P*	There is an orange. INCOMPATIBLE / There is not an orange.
(4.2)	*P*; NOT *P* INCOMPATIBLE	There is an orange; There is not an orange. / INCOMPATIBLE
(5.1)**	All *P* are *Q*. All *Q* are *R*. ǂ All *P* are *R*.	All felines are cats. All cats are finicky. ǂ ∴ All felines are finicky.
(5.2)**	All *P* are *Q*. *R* is *P*. ǂ *R* is *Q*.	All cats are finicky. Fritz is a cat. ǂ ∴ Fritz is finicky.
(6.1)**	*P* is *Q*. Something is *Q*.	Fritz is finicky. ∴ Something is finicky.
(6.2)**	For all *P*, $P \rightarrow Q$. ǂ Something is *P*. *R* is *P*. ǂ If *R* is *P*, then *R* is *Q*. ǂ For *R*, *Q*.	All citrus fruits are bitter. ǂ Something is a citrus fruit. A lemon is a citrus fruit. ǂ If a lemon is a citrus fruit, then it is bitter. ǂ ∴ A lemon is bitter.

Fig. 11: The inference schemas of mental logic.

Fritz, and P corresponds to the property "is finicky". Formally,

$$(\forall x)\text{isfinicky}(x) \Rightarrow \text{isfinicky}(\text{Fritz})$$

Clearly, the definition of a domain of discourse made the rule less trivial, but by doing so it spoiled what in logic is *the* quality of logic, to wit, sheer generality. This is to say that by defining a domain of discourse that has to do with the real world, the rules lose their purely formal character and become subject to the many vagaries of natural language. In the case above, the problem begins precisely with the domain of discourse, *cats*, and it is only natural that it should be so if we have in mind what was said above about categorization. That is, what exactly are or count as cats: domestic cats alone?; also feral ones?; all felines?; cartoon cats? Is the Cheshire Cat a cat? And then we get to the property in question, "is finicky", and we are doubling the problem, for what it is to be finicky is certainly not a simple thing to say. Moreover, one may, consciously or unconsciously, disagree that all cats are finicky; one certainly has one's own ideas about cats. Even if we agree about what *is finicky* means, and agree to what *cats* are supposed to be and that they are supposed to be finicky, in an experiment on semantic memory we still predict widely varying reaction times to answers of True or False to the statement "This cat is finicky" when the subject is shown a picture displaying

- a relaxed domestic cat
- a starving-looking, hunting cat
- the Cheshire Cat
- a roaring lion
- a cubist drawing of a cat

The predictable results of such an experiment would involve a plethora of psychological phenomena ranging from implicit association biases to Stroop-like inhibitory effects, all belying the triviality of the inference rule of universal elimination.

The inference schemas for existential quantification can be shown to involve even more complex psychological phenomena. *Existential introduction*, or *existential generalization*, the inference from a proposition stating that a particular constant c is or has the property P to the affirmation that someone or something is or has the property P. Formally, $Pc \Rightarrow \exists x Px$. Resorting to Fritz again,

$$\text{isfinicky}(\text{Fritz}) \Rightarrow (\exists x)\text{isfinicky}$$

The triviality of this rule or schema of inference is, if anything, even more obvious than that of the rule or schema for universal elimination, but, again, this is only so if one is contented with using the rule in its contentless generality; once you put some flesh to the logical bones, triviality vanishes. Without entering into the logico-metaphysical perplexities of existential import, one may accept that Fritz is a *felus catus* without for that being willing to accept the statement that something is a *felus catus*. The reason might be that one interprets the copula (i.e., the verb *be*) in metaphysically distinct ways in "Fritz is a *felus catus*" and "Something is a *felus catus*"; for instance, while the first statement may not appear to entail any existential import, this import might appear to be entailed in the second statement: due, say, to ignorance with respect to Latin, this might appear inappropriate.

Another illustration involving different reasons could be obtained with the statements "God is omnipotent" and "Someone is omnipotent": that one accepts that God is omnipotent, that does not entail that one accepts that someone or something is omnipotent. Kant would very likely be one of them, as, according to him, while one cannot accept that God is without accepting that he is omnipotent (the concept *omnipotence* makes part of the concept *God*, in the same way that one cannot say of a triangle that it does not have three sides), one can indeed reject both concepts, without for that entailing a contradiction (see Kant's *Critique of Pure Reason*, A 592ff/B 620ff); that is to say that I can conceive that, if there is God, then he is omnipotent, omnipotence being an attribute of God alone, but I can also conceive that there is no God, and that therefore no one is omnipotent.

Thus, whereas the logical triviality of universal elimination may vanish in face of psychological phenomena related to perception and memory, i.e., the thought processes of categorization and association, the triviality of existential introduction may irritate higher thought processes involving, for instance, religious beliefs, social taboos, superstitions, etc, or just the plain attitudes and stereotypes well-known to social psychology. For instance, one may, somehow confusedly, accept that a certain person is P or has property P without for that accepting that someone – i.e., at least one person – is P or has property P (i.e., that there is such a thing as P[42]). This psychological phenomenon, *denial*, thus disproves the claimed universally trivial character of the inference schemas of ML.

This shows how the *psychological* restrictions that actual thought imposes on the schemas for quantification are far more complex and numerous that those that logic itself contemplates. I was saving these for last, for, though the psychological restrictions offer counterexamples, and thus contradict the claims to universality, they, of course, do not really contradict the supposedly basic, elementary character of the schemas and rules of ML and ND alike. These *logical* restrictions are as follows: (1) in the case of universal elimination, it is required that the instantiation be produced by uniformly replacing each occurrence of the bound variable by a constant; (2) as for existential generalization, slightly more complex, it is necessary that (a) the variable to be used in the generalization does not occur in the sentence that is generalized, and (b) the generalization is obtained by replacing at least one occurrence of the constant by the variable, no other changes being made.

Regarding (1), the problem lies in the adverb of manner *uniformly*: this uniform replacement of each occurrence of the bound variable (say, x in $\forall xPx$) by a constant (c) means simply that in order to go from $\forall xPx$ to Pc – i.e., to eliminate or instantiate the universal quantifier – one or more propositional inferences are required.[43] Resorting again to the domain of cats and to Fritz, the universal instantiation actually runs as follows: "All cats are finicky. Fritz is a cat. Therefore Fritz is finicky." Let Fx mean "x is finicky" and Kx mean "x is a cat"; then, what is involved in the instance featuring Fritz is the argument[44]

$$\forall x[(Kx \rightarrow Fx) \wedge Kc] \vdash Fc$$

This can hardly be said basic or elementary, at least if by elementary one means indecomposable – or as close to that as possible. As for (2), restriction (a) means,

[42] In the sense that, for instance, there is wealth when there are wealthy people, just as there is the law when there are laws; see, for instance, James (1907/1970), p. 158.

[43] This restriction is often formulated in the following terms: if c is a variable, then it must not already be quantified anywhere in Px.

[44] Read as "For all the x, if x is a cat, then x is finicky; (and) c is a cat. Therefore, c is finicky."

more rigorously, that x cannot appear as a free variable in Pc;[45] less rigorously, that if there is one element belonging to the universe of discourse that is P or has property P (here: c), then we know (i.e., it follows) that something exists in the defined universe that is P or has property P. Rule (b) states the fact that from the assumption that Fritz, our c, is a cat and Fritz is finicky (formally: $Kc \wedge Fc$), it follows without doubt that at least something is a cat and Fritz is finicky ($\exists x(Kx \wedge Fc)$) and also that Fritz is a cat and something is finicky ($\exists x(Kc \wedge Fx)$).

As announced, the rules or schemas for *universal introduction* or *generalization* and for *existential elimination* or *instantiation*,[46] are significantly more complex, with more restrictions, not being simply the inverse of those for universal elimination and existential introduction, thus contradicting the *inversion principle* of ND that lies at its very foundations (see Prawitz, 1965/2006, p. 32ff). For instance, while one is tempted to see the former pair of rules in the light of the triviality that characterizes the latter, two instances suffice to show how wrong one is: from a statement such as "Something is finicky", one obviously is not allowed to infer that, say, "A pear is finicky", though the rule, representable symbolically as $\exists x Px \Rightarrow Pc$, in its simplicity appears to allow so; as for the rule of inference for universal introduction, it is obvious that from the fact that Fritz is finicky, one cannot infer that everything is finicky, thus belying the simplicity of the rule $Pc \Rightarrow \forall x Px$.

Let us linger for a moment on universal introduction. The idea behind this rule of inference is that if one can prove something regarding an individual, say c, without any assumptions that distinguish c from any other individual x, then what one proves of c can be proved of anything whatsoever, i.e., it holds for absolutely everything. This might be obvious, but we saw that from the statement "Fritz is finicky" one can clearly not infer that, say, "A pear is finicky"; there must be one or more restrictions behind this façade of triviality and simplicity. This is indeed the case, with the obvious restriction being that c cannot be identified as a specific element of the universe of discourse. Another restriction is that x must not appear as a free variable in Pc – or, in other words, x cannot appear in the sentence to be generalized. These requirements should appear more 'reasonable' with a proof; let us use the following argument:

All fish are swimming animals.
All swimming animals are pasta-eating animals.
∴ All fish are pasta-eating animals.

Let Fx mean "x is a fish", Sx mean "x is a swimming animal", and Px mean "x is a pasta-eating animal"; then, the argument above can be represented as

$$\forall x(Fx \to Sx), \forall x(Sx \to Px) \vdash \forall x(Fx \to Px)$$

and the proof for its form can be constructed as

[45] In predicate or first-order logic, a variable is free if it falls outside the scope of a quantifier; otherwise, it is bound. For instance, in the expression $\exists x(Fx \wedge Ky)$, x is a bound variable, and y is a free variable.

[46] Note that in this case existential elimination is not necessarily a synonym for existential instantiation, as they may operate differently; unless otherwise stated, assume we are talking of the former.

1	$\forall x(Fx \rightarrow Sx)$	Assumption
2	$\forall x(Sx \rightarrow Px)$	Assumption
3	$Fc \rightarrow Sc$	1 $\forall E$
4	$Sc \rightarrow Pc$	2 $\forall E$
5	$Fc \rightarrow Pc$	3, 4 Hypothetical syllogism
6	$\forall x(Fx \rightarrow Px)$	5 $\forall I$

Note how c, which designates an individual about which no assumption was made (i.e., it does not feature in steps 1 and 2 of the proof above), is introduced only in steps 3 and 4 by the means of $\forall E$; this makes it representative of all individuals, whatever is proved of it holding thus for any individual whatsoever. At 5, we prove that if c is a fish, then c is a pasta-eating animal, and as the proof is general, because we made no assumptions about c, we could have also proved that, say, if d is a fish, then d is a pasta-eating animal, and so on for any individual. What we have done at 6 by applying $\forall I$ is, we have proved that if x is a fish, then x is a pasta-eating animal. If, on the contrary, we begin by making an assumption about c, namely that c is supposed to represent a specific element of the universe of discourse that has the property P, then we are not allowed to generalize it by applying $\forall I$. Then, oddly enough, the argument $Fc \vdash \forall xFx$, equivalent for our purposes (see above) to $Fc \Rightarrow \forall xFx$, our rule of inference for universal generalization, is actually invalid, as the proof shows:

1	Fc	Assumption
2	$\forall xFx$	1 $\forall I$ (incorrect)

And we already know why: because from the assumption that Fritz is finicky it does not hold that everything is finicky.

If the reader thinks that this is wildly remote from what could be seen as a basic thought process, then a look at the restrictions imposed on existential elimination, representable by the rule of inference $\exists xPx \Rightarrow Pc$, will only strengthen this opinion: (a) each occurrence of the bound variable is uniformly replaced by a constant, no other changes being made, and (b) the constant does not appear in any earlier premise of the argument. This all reduces to saying that if you have been able to prove that there is some, unspecified thing, that is P or has property P, then you might well hypothesize that c is P or has property P; discharge the hypothesis and assert that indeed there is a c that is P or has property P: the 'strength' alone of $\exists xPx$ allows you to do this. This is what existential elimination is all about. Now the problem is, of course, how one proves anything regarding an unspecified thing. One cannot; all one can do is to prove something regarding a universe of discourse, and then instantiate, or eliminate the existential quantifier. For instance, suppose we have proved that there is at least *something* that is both finicky and messy ($\exists x(Fx \wedge Mx)$); we can then hypothesize that c has one of these properties; the hypothesis indeed leads us to the conclusion that there is at least one individual that has that property. The proof runs like this:

1	$\exists x(Fx \wedge Mx)$	Assumption
2	$\mid Fc \wedge Mc$	Hypothesis (for $\exists E$)
3	$\mid Fc$	2 $\wedge E$
4	$\mid \exists xFx$	3 $\exists I$
5	$\exists xFx$	1, 2-4 $\exists E$

I might be wrong, but this sounds very much like circularity disguised, again, by the need to add and/or suppress further premises to the alluring simplicity of the rule of inference as it is stated in formal terms.

Our problems with the schemas for quantificational inference are not yet over. Let us revisit the rule of inference for existential introduction: besides the restrictions specified above – let us call them *logical restrictions* – it also imposes *metaphysical restrictions*, which are actually *metaphysical presuppositions*: (1) all the proper names, our cs, must refer to actually existing individuals, and (2) at least one individual exists. How (1) simply goes against, for instance, religious thought, it is hardly necessary to illustrate, and we can only suspect that either religious thinking does not involve basic, universal thought processes, or these are actually not basic and/or universal at all. If one sticks to the latter, i.e., the schema for existential introduction is basic and universal, then one is making it indeed unthinkable for a vast number of otherwise rational people who *cannot* think, given a, say, mythological being such as Zeus (i.e., Mz meaning "z is a mythological being"), that it is *not* really the case that a mythological being exists ($\exists xMx$).[47] As for (2), it takes training in metaphysical matters in order to presuppose that at least one thing exists (in other words: to presuppose existence) when making an existential introduction inference, and this training is not meant to assure the thinker of the above, that is, that at least one individual exists.

As I have already proved my point, I will leave it to the curiosity and the ingenuity of the reader to find examples that contradict the claims of the basic, elementary character of these and of the propositional inference rules of ND and ML made by both logicians and many philosophers alike. I invite the reader to further pursue this exercise in the extensions of classical logic (see Appendix, Fig. 13), the approach to which is rather easy once one masters the basics of classical deductive logic.[48]

[47] This problem is eliminated in free logic, an extension of classical logic: while the quantifiers are interpreted in the classical way, its singular terms may be wholly non-denoting or may denote objects outside the universe of discourse. See Lambert (1967) for a smooth introduction to free logic.

[48] I particularly recommend an approach of epistemic logic with relation to thinkability. For instance, the epistemic logic systems known as **S4** and **S5** are characterized by the axioms of positive and negative introspection, respectively, expressing the facts that if one knows a proposition p, then one knows the proposition p (formally: $Kp \rightarrow KKp$) and that if one does not know p, then one knows that one does not know p ($\neg Kp \rightarrow K\neg Kp$). Epistemic logic is an example par excellence of how problematic it can be to pretend to see logic as descriptive, let alone normative, of human thought, given that it claims to describe the epistemic behavior of rational beings. As a matter of fact, the workings of epistemic logic are such that it is plagued by the problem of *logical omniscience*, a wholly unrealistic property in the case of humans and highly undesirable in an AI context. While some recognize that epistemic logic is a logic of, or for, ideal reasoners (e.g., Rescher, 1974), namely machine 'reasoners,' others see it as also a logic of human reasoning (e.g., van Benthem, 2006). The fact that it deals with thought processes conducive to knowledge makes it even more problematic (see for instance Hocutt, 1972, for a challenging of epistemic logic both as a logic and as epistemic).

III. Conclusions

While being incapable of a clear-cut use of the terms *thought* and *logic*, we often hastily judge instances of thought as *illogical*, whereas we tend to look upon logic as actually *thinkable*. Caution is advised in both cases, as what illogical thought might be depends on a concept of logic as a discipline, and the decision whether logic actually is thinkable requires an empirical study; the inverse is also true, that is, what logical thought is depends on a concept of logic, and what unthinkable logic might be depends, too, on observation of the ways people do indeed think.

Regarding *(il)logicality*, we have to take into consideration the progress (some would say degeneration) logic and its allied set theory have undergone since the early decades of the 20th century, a progress that is of particular interest to psychology as it was more often than not motivated by the psychological limitations of humans. This progress has occurred in the direction of an elaboration of logics and extensions to set theory as *theories of perception*. We thus see our psychological limitations justified as logical in nature: it is not illogical to accept (true) contradictions or to reject the PEM if one is thinking in the eyes of, say, paraconsistent and intuitionistic logics, respectively; according to relevance logics, one cannot be said illogical when disregarding argument form in light of additional, relevant information; and many-, sometimes even infinite-valued, logics permit us to use concepts and manipulate propositions allowing for the uncertainty and vagueness that characterizes natural languages. An interesting aspect of the acceptance of the above as actually logical thinking is its application in AI: we often first had to simulate machine applications of logical systems to realize that classical logic alone fell short of describing human successfull thought processes. This is particularly true of theories of human semantic memory, but it also bears on issues such as belief revision, limited availability of information, and even metaphysical considerations have been given room in the *new* logics. But although many of these most well-known new logics have been around for quite a while, and despite the fact that they were motivated by psychological and empirical constraints, we still have not put them into relation with psychological issues. There is yet no telling how much light these new logics might shed on problematic issues involving clinical conditions of thought disorder such as schizophrenia, but one thing is certain: we can no longer approach them from a point of view of illogicality established a priori, in the spirit of *old* classical logic (where *old* here means simply preceding in time).

This is so not only because this logic falls short of accounting for human reasoning limitations that nevertheless do not essentially attempt against logicality, but also because it has seen its claims to *thinkability*, supported by claims of universality, sheer generality, and even triviality, challenged. In fact, an analysis of the inference schemas of the so-called mental logic and of natural deduction, where both terms *mental* and *natural* are meant to convey the fact that they reflect actual ways of thinking, shows that none of the above features claimed by classical logic stand the proof of actual thinking: firstly, because they are very remote indeed from the basic character attributed to them, involving complex reasoning operations that are not at all obvious; secondly, because psychological and even cultural restraints produce counterexamples to their claimed generality and universality. That is to say that classical logic – and some of its extensions (in particular epistemic logic) – makes for what in fact can be seen as *unthinkable logic*.

It is true that much academic work has been done into the relations between classical logic and the psychology of reasoning, but all remains practically to be done in the investigations of the intersections between the new logics and the recent extensions of set theory and all levels of thinking. The work done so far has been strongly limited by both focusing on classical logic and concentrating solely on higher levels of thinking, commonly labeled as reasoning. But the new logics and the recent extensions to set theory have progressed in the direction of theories of perception, thus tackling all levels of thinking, and promising therefore to provide us with a more encompassing study of the (il)logicality of thought, while at the same time favoring a study of the (un)thinkability of logic(s).

Acknowledgments: I wish to thank the Portuguese Foundation for Science and Technology (FCT) and the European Social Fund for the post-doctoral fellowship that made this work, among others, possible.

Appendix: Further Logical Issues

Laws	Algebra of Propositions		Algebra of Sets	
Idempotent laws	$P \vee P \equiv P$	$P \wedge P \equiv P$	$A \cup A = A$	$A \cap A = A$
Associative laws	$(P \vee Q) \vee R \equiv$ $P \vee (Q \vee R)$	$(P \wedge Q) \wedge R \equiv$ $P \wedge (Q \wedge R)$	$(A \cup B) \cup C =$ $A \cup (B \cup C)$	$(A \cap B) \cap C =$ $A \cap (B \cap C)$
Commutative laws	$P \vee Q \equiv Q \vee P$	$P \wedge Q \equiv Q \wedge P$	$A \cup B = B \cup A$	$A \cap B = B \cap A$
Distributive laws	$P \vee (Q \wedge R) \equiv$ $(P \vee Q) \wedge (P \vee R)$	$P \wedge (Q \vee R) \equiv$ $(P \wedge Q) \vee (P \wedge R)$	$A \cup (B \cap C) =$ $(A \cup B) \cap (A \cup C)$	$A \cap (B \cup C) =$ $(A \cap B) \cup (A \cap C)$
Identity laws	$P \vee \mathrm{T} \equiv P$ $P \vee \mathrm{T} \equiv \mathrm{T}$	$P \wedge \mathrm{F} \equiv P$ $P \wedge \mathrm{F} \equiv \mathrm{F}$	$A \cup \emptyset = A$ $A \cup \mathbf{U} = \mathbf{U}$	$A \cap \mathbf{U} = A$ $A \cap \emptyset = \emptyset$
Involution law	${\sim}{\sim}P \equiv P$		$(A^c)^c = A$	
Complement laws	$P \vee {\sim}P \equiv \mathrm{T}$ $P \wedge {\sim}P \equiv \mathrm{F}$	${\sim}\mathrm{T} \equiv \mathrm{F}$ ${\sim}\mathrm{F} \equiv \mathrm{T}$	$A \cup A^c = \mathbf{U}$ $\mathbf{U}^c = \emptyset$	$A \cap A^c = \emptyset$ $\emptyset^c = \mathbf{U}$
De Morgan's laws	${\sim}(P \vee Q) \equiv {\sim}P \wedge {\sim}Q$	${\sim}(P \wedge Q) \equiv {\sim}P \vee {\sim}Q$	$(A \cup B)^c = A^c \cap B^c$	$(A \cap B)^c = A^c \cup B^c$

Fig. 12: 'Similarity' between the laws of the algebras of propositions and of sets.

Classical Logic

First-Order Predicate Calculus

[& 'Classical' Set Theory; also Types Theory, and Categories Theory]

Non-Classical Logics

[& 'Non-Classical' Set Theory: Fuzzy Set Theory; Rough Set Theory]

Extended Logics (Extensions of Classical Logic)

Logics	Main extending properties
Modal logics	Non-truth functional operators: *necessarily* (□); *possibly* (◊)
• Epistemic logics	Epistemic operators: *it is known* (*K*); *it is believed* (*B*)
• Deontic logics	Deontic operators: *it is obligatory* (*O*); *it is permitted* (*P*)
• Temporal or tense logics	'Temporal' operators: *future* (*G*); *past* (*H*); etc.
Intensional logics	Additional quantifiers for *intensions*
Inductive logic	Degrees of strength (vs. logical validity)
Free logic	Allows for non-denoting singular terms (e.g., Zeus; the present king of France; etc)

Deviant Logics ('Rivals' of Classical Logic)

Logics	Main deviant properties
Intuitionistic logic	Rejection of PEM and of double negative elimination
• Minimal logic	Besides the law of PEM, also rejects the principle of explosion (*ex contradictione quodlibet,* from a contradiction follows anything)
Many-valued logics	Rejection of the principle of bivalence
• 3-valued logics of Kleene and Łukasiewicz	Three truth-values: *true* (1); *false* (0); *neither true nor false, i.e. indeterminate*(*i*) or *undeterminable* (*u*)
• Fuzzy logics	Any truth-value ranging from 0 to 1 (e.g. 0,3)
Quantum logics	Rejection of distributive laws
Paraconsistent logics	Rejection of the principles of explosion (*ex contradictione quodlibet*) and (sometimes) of PC
Relevance logics	Rejection of monotonicity of entailment (in *A* → *B*, *A* and *B* must be relevantly related)

Fig. 13: Some of the most well-known non-classical logics.

	Introduction Rules	Elimination Rules	
(1.1)	$\dfrac{A \quad B}{A \wedge B}\wedge I$	$\dfrac{A \wedge B}{A}\wedge E \qquad \dfrac{A \wedge B}{B}\wedge E$	(1.2)
(2.1)	$\dfrac{A \vee B}{A}\vee I \qquad \dfrac{A \vee B}{B}\vee I$	$\dfrac{A \vee B \quad \overset{(A)}{\overline{C}} \quad \overset{(B)}{\overline{C}}}{C}\vee E$	(2.2)
(3.1)	$\dfrac{\overset{(A)}{\overline{B}}}{A \rightarrow B}\rightarrow I$	$\dfrac{A \quad A \rightarrow B}{B}\rightarrow E$	(3.2)
(4.1)	$\dfrac{\overset{(A)}{\overline{\curlywedge}}}{\sim A}\sim I$	$\dfrac{A \quad \sim A}{\curlywedge}\sim E$	(4.2)
(5.1)	$\dfrac{A}{\forall x A^a_x}\forall I$	$\dfrac{\forall x A}{A^x_t}\forall E$	(5.2)
(6.1)	$\dfrac{A^x_t}{\exists x A}\exists I$	$\dfrac{\exists x A \quad \overset{A^x_a}{\overline{B}}}{B}\exists E$	(6.2)

Fig. 14: The inference rules of natural deduction.

References

Akiskal, H. S. (2008). The Mental Status Examination. In S. H. Fatemi & P. J. Clayton (eds.), *The Medical Basis of Psychiatry* (pp. 3-16). 3rd ed. Totowa, NJ: Humana Press.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., revised). Washington, DC: Author.

Aristotle (DI/1941). *Categoriae* [*Categories*]. Trans. E. M. Edgehill, in R. McKeon (ed.), *The Basic Works of Aristotle* (pp. 7-39). New York: Random House.

Aristotle (DI/1941). *De Interpretatione* [*On Interpretation*]. Trans. E. M. Edgehill, in R. McKeon (ed.), *The Basic Works of Aristotle* (pp. 40-64). New York: Random House.

Augusto, L. M. (2011). Food for thought: Rethinking thinking. [forthcoming]

Blanchette, P. A. (2001). Logical consequence. In L. Goble (ed.), *The Blackwell Guide to Philosophical Logic* (pp. 115-135). Oxford: Blackwell.

Bleuler, E. (1911/1950). *Dementia Praecox, or the Group of the Schizophrenias.* Transl. by J. Zinkin, New York: International Universities Press.

Boole, G. (1847/1948). *The Mathematical Analysis of Logic*. New York: The Philosophical Library.

Boole, G. (1854/1958). *An Investigation of the Laws of Thought*. New York: Dover.

Braine, M. D. S. (1990). The "Natural Logic" approach to reasoning. In W. F. Overton (ed.), *Reasoning, Necessity, and Logic: Developmental Perspectives* (pp. 133-157). Hillsdale, NJ: Lawrence Erlbaum Associates.

Braine, M. D. S. (1998). Steps toward a mental-predicate logic. In M. D. S. Braine & D. P. O'Brien (eds.), *Mental Logic* (pp. 273-331). Mahwah, NJ: Lawrence Erlbaum Associates.

Braine, M. D. S., & O'Brien, D. P. (eds.) (1998). *Mental Logic*. Mahwah, NJ: Lawrence Erlbaum Associates.

Brouwer, L. E. J. (1907/1975). On the Foundations of Mathematics. In A. Heyting (ed.), *Collected Works I: Philosophy and Foundations of Mathematics* (pp. 11–101). Amsterdam: North-Holland.

Brouwer, L. E. J. (1908/1975). The Unreliability of the Logical Principles. In A. Heyting (ed.), *Collected Works I: Philosophy and Foundations of Mathematics* (pp. 107–111). Amsterdam: North-Holland.

Burge, T. (2010). *Origins of Objectivity*. Oxford: Oxford University Press.

Bürgy, M. (2008). The concept of psychosis: Historical and phenomenological aspects. *Schizophrenia Bulletin, 34*, 1200-1210.

Cantor, G. (1895). Beiträge zur Begründung der transfiniten Mengenlehre, I. [Contributions to the foundation of transfinite set theory, I]. *Mathematische Annalen, 46*, 481-512.

Chaika, E. O. (1974). A linguist looks at 'schizophrenic' language. *Brain and Language, 1*, 257-276.

Chen, E. Y., Wilkins, A. J., & McKenna, P. J. (1994). Semantic memory is both impaired and anomalous in schizophrenia. *Psychological Medicine, 24*, 193-202.

Collins, A. M. & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, 82, 407-428.

Collins, A. M. & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 240-248.

Ebbinghaus, H. (1885/1913). *Memory: A Contribution to Experimental Psychology*. Trans. by H. A. Ruger & C. E. Bussenius, New York: Teachers College, Columbia University.

Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. In D. Touretzky (ed.), *Connectionist Approaches to Language Learning* (pp. 91-122). Dordrecht: Kluwer.

Elvevåg, B., Weickert, T., Wechsler, M., Coppola, R., Weinberger, D. R., & Goldberg, T. E. (2002). An investigation of the integrity of semantic boundaries in schizophrenia. *Schizophrenia Research*, *53*, 187–198.

Elvevåg, B., Heit, E., Storms, G., & Goldberg, T. (2005). Category content and structure in schizophrenia: An evaluation using the instantiation principle. *Neuropsychology, 19*, 371-380.

Erdmann, B. (1892). *Logische Elementarlehre* [*Elementary doctrine of logic*]. Halle: Niemeyer.

Evans, J. S. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin, 128*, 978-996.

Ferreirós, J. (2000). *The Labyrinth of Thought: A History of Set Theory and Its Role in Modern Mathematics*. 2nd revised ed. Basel – Boston – Berlin: Birkhäuser.

Fiddick, L., Cosmides, L., & Tooby, J. (2000). No interpretation without representation: The role of domain-specific representations and inferences in the Wason selection task. *Cognition, 75*, 1-79.

Fodor, J. (1975). *The Language of Thought*. Cambridge, Massachusetts: Harvard University Press.

Fodor, J., & Pylyshyn, Z. (1988). Connectionism and Cognitive Architecture: a Critical Analysis. *Cognition, 28*, 3–71.

Frege, G. (1884/1953). *The Foundations of Arithmetic: A Logico-mathematical Inquiry into the Concept of Number*. Transl. by J. L. Austin. 2nd ed. Evanston, Ill.: Northwestern University Press.

Gentzen, G. (1934-5/1969). Investigations into logical deduction. In M. E. Szabo (ed.), *The Collected Papers of Gerhard Gentzen* (pp. 68-131). Amsterdam: North-Holland.

Gigerenzer, G. (2000). *Adaptive Thinking: Rationality in the Real World*. Oxford: Oxford University Press.

Goldberg, T. E. & Weinberger, D. R. (2000). Thought disorder in schizophrenia: A reappraisal of some older formulations and an overview of some recent studies. *Cognitive Neuropsychiatry, 5*, 1-19.

Haack, S. (1996). *Deviant Logic, Fuzzy Logic. Beyond the Formalism*. Chicago and London: The University of Chicago Press.

Hanna, R. (2006). *Rationality and Logic*. Cambridge, Massachusetts: MIT Press.

Hilbert, D. (1931/1996). The grounding of elementary number theory. In W. B. Ewald (ed.), *From Kant to Hilbert: A Source Book in the Foundations of Mathematics* (pp. 1148-56). 2 vols. Oxford: Oxford University Press.

Hocutt, M. O. (1972). Is epistemic logic possible? *Notre Dame Journal of Formal Logic, 13*, 433-453.

Hume, D. (1748/1999). *An Enquiry Concerning Human Understanding*. Ed. T. L. Beauchamp. Oxford: Oxford University Press.

Husserl, E. (1900/1970). *Logical Investigations*. (2 vols.) Transl. by J. N. Findlay. London: Routledge and Kegan Paul.

James, W. (1907/1970). *Pragmatism*. Cleveland & New York: Meridian Books.

Jaśkowski, S. (1934). On the rules of suppositions in formal logic. *Studia Logica, 1*, 5-32.

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

Jung, C. G. (1904-10/1973). *Studies in Word Association*. In *The Collected Works of C. G. Jung*. Vol. 2: *Experimental Researches*. Transl. by G. Adler & R. F. C. Hull, Princeton, NJ: Princeton University Press.

Jung, C. G. (1907/1960). *The Psychology of Dementia Praecox*. In *The Collected Works of C. G. Jung*. Vol. 3: *The Psychogenesis of Mental Disease* (pp. 1-151). Transl. by R. F. C. Hull, Princeton, NJ: Princeton University Press.

Kant, I. (1781; 1787/2003). *Critique of Pure Reason*. Trans. by N. K. Smith. Hampshire: Palgrave MacMillan.

Kant, I. (1800/1885). *Kant's Introduction to Logic*. Trans. by T. K. Abbott. London: Longmans, Green, & Co.

Kneale, W., & Kneale, M. (1962). *The Development of Logic*. Oxford: Clarendon Press.

Kraepelin, E. (1896/1919). *Dementia Praecox and Paraphrenia*. Trans. by R. M. Barclay & G. M. Robertson. Edinburgh: Livingstone [extracted from: Kraepelin, E. (18913). *Psychiatrie*. 8th ed., vol. III. Leipzig: Barth].

Kusch, M. (1995). *Psychologism: A Case Study in the Sociology of Philosophical Knowledge*. London: Routledge.

Lambert, K. (1967). Free logic and the concept of existence. *Notre Dame Journal of Formal Logic, 8*, 133-144.

Leeson, V. C., Simpson, A., McKenna, P. J., & Laws, K. R. (2005). Executive inhibition and semantic association in schizophrenia. *Schizophrenia Research, 74*, 61-67.

Locke, J. (1690/1959). *An Essay concerning Human Understanding* (2 Vols.). New York: Dover Publications.

Luria, A. R. (1976). *Cognitive Development: Its Social and Cultural Foundations*. Cambridge, MA: Harvard University Press.

Maher, B. A. (1983). A tentative theory of schizophrenic utterance. In B. A. Maher & W. B. Maher (eds.), *Progress in Experimental Personality Research*, vol. 12: *Psychopathology* (pp. 1-52). New York: Academic Press.

Margolis, E. & Laurence, S. (eds.) (1999). *Concepts: Core Readings*. Cambridge, MA: MIT Press.

McKay, A. P., McKenna, P. J., Bentham, P., Mortimer, A. M., Holbery, A., & Hodges, J. R. (1996). Semantic memory is impaired in schizophrenia. *Biological Psychiatry, 39*, 929–937.

McKenna, P., & Oh, T. (2005). *Schizophrenic Speech. Making Sense of Bathroots and Ponds that Fall in Doorways*. Cambridge: Cambridge University Press.

Medin, D. L. & Rips, L. J. (2005). Concepts and categories: Memory, meaning, and metaphysics. K. J. Holyoak & R. G. Morrison (eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 37-72), Cambridge: Cambridge University Press.

Mill, J. (1829). *Analysis of the Phenomena of the Human Mind*. London: Baldwin and Cradock.

Mill, J. S. (1843). *A System of Logic Ratiocinative and Inductive: Being a connected view of the Principles of Evidence and the Methods of Scientific Investigation*. In *The Collected Works of John Stuart Mill*. Vol. 7 (Books I-III), 8 (Books IV-VI). Ed. By J. M. Robson. London and Toronto: Routledge and University of Toronto Press, 1973-4.

Molczanow, A. (2002). Quantification and inference. *The Monist, 85*, 144-159.

Molczanow, A. (2004). *Logical Aspects of Quantification in Natural Language*. Bloomington, Indiana: Unlimited Publishing.

Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

Niznikiewicz, M., Mittal, M. S., Nestor, P. G., & McCarley, R. W. (2010). Abnormal inhibitory processes in semantic networks in schizophrenia. *International Journal of Psychophysiology, 75*, 133-140.

Pagliani, P. & Chakraborty, M. (2008). *A Geometry of Approximation: Rough Set Theory: Logic, Algebras and Topology of Conceptual Patterns*. Springer.

Pawlak, Z. (1982). Rough sets. *International Journal of Parallel Programming, 11*, 341–356.

Plato (DI/1961), *Euthyphro*. Trans. by L. Cooper. In E. Hamilton & H. Cairns (eds.), *Plato. The Collected Dialogues, Including the Letters* (pp. 170-185). Ithaca: Princeton University Press.

Prawitz, D. (1965/2006). *Natural Deduction: A Proof-Theoretical Study*. Mineola, NY: Dover.

Priest, G. (2008). *An Introduction to Non-Classical Logic*. Cambridge: Cambridge University Press.

Quillian, M. R. (1969). The Teachable Language Comprehender: A simulation program and theory of language. *Communications of the ACM*, 12, 459-476.

Quine, W. V. O. (1970). *Philosophy of Logic*. Englewood Cliffs, N.J.: Prentice-Hall.

Rescher, N. (1974). On alternatives in epistemic logic. In N. Rescher, *Studies in Modality. American Philosophical Quarterly* (pp. 99-114). Monograph Series no. 8, Oxford: Basil Blackwell.

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

Rumelhart, D. & McClelland, J. (and the PDP Research Group) (1986). *Parallel Distributed Processing*. Vol. I. Cambridge, Massachusetts: MIT Press.

Saussure, F. de (1916/1977). *Course in General Linguistics*. Transl. by W. Baskin, Glasgow: Fontana Collins.

Schreber, D. P. (1903/2000). *Memoirs of My Nervous Illness*. Ed. and transl. by I. Macalpine & R. A. Hunter, New York: New York Review of Books.

Scribner, S. (1997). *Mind and Social Practice*. Oxford: Oxford University Press.

Smiley, T. (1960). Sense without denotation. *Analysis, 20*, 125-135.

Smithson, M. & Oden, G. C. (1999). Fuzzy set theory and applications in psychology. In H.-J. Zimmermann (ed.), *Practical Applications of Fuzzy Technologies* (pp. 557-588), Boston – London – Dordrecht: Kluwer.

Smithson, M. & Verkuilen, J. (2006). *Fuzzy Set Theory: Applications in the Social Sciences*. Thousand Oaks – London – New Delhi: Sage Publications.

Stenning, K., & van Lambalgen, M. (2008). *Human Reasoning and Cognitive Science*. Cambridge, MA: MIT Press.

Storms, G., & Elvevåg, B. (2010). Thinking about semantic concepts in schizophrenia: the more familiar the less deviation. *Schizophrenia Research, 116*, 295-296.

Tamlyn, D., McKenna, P. J., Mortimer, A. M., Lund, C. E., Hammond, S., & Baddeley, A. D. (1992). Memory impairment in schizophrenia: Its extent, affiliations and neuropsychological character. *Psychological Medicine, 22*, 101–115.

van Benthem, J. (2006). Epistemic logic and epistemology: The state of their affairs. *Philosophical Studies, 128*, 49-76.

Wason, P. (1966). Reasoning. In B. M. Foss (Ed.), *New Horizons in Psychology* (pp. 135-151). Harmondsworth: Penguin.

Wason, P. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20*, 273-281.

Waters, F. A. V., Badcock, J. C., Maybery, M. T., & Michie, P. T. (2003). Inhibition in schizophrenia: association with auditory hallucinations. *Schizophrenia Research, 62*, 275– 280.

Wernick, W. (1942). Complete sets of logical functions. *Transactions of the American Mathematical Society, 51*, 117-132.

World Health Organization (2007). *International Classification of Diseases and Related Health Problems* (10th revision). Geneva: Author. Online edition: http://apps.who.int/classifications/apps/icd/icd10online/

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*, 338-353.