Vol. 5 (2013) 13-21
University of Reading
ISSN 2040-3461
LANGUAGE STUDIES WORKING PAPERS
Editors: C. Ciarlo and D.S. Giannoni

Influence of World Knowledge and Context on the Comprehension of Natural Language Translation of Logical Formulas

Luca Cilibrasi and Matteo Pascucci

In this paper we present an approach to conditional reasoning tasks based on two main ideas. The first idea is that, in contrast with what is usually assumed, an 'if… then…' sentence is not an adequate translation in natural language of a logical formula containing a material implication as its principal operator. The second idea is that when subjects are required to check the validity of a sentence in a task, their inferences are not driven uniquely by the content of the sentence, but also by other information embedded in the task scenario and, eventually, by their knowledge about the topic (i.e. information stored in memory). Data from 30 subjects tested on six different tasks are reported as evidence for our approach. The results show that conditional tasks are significantly more complex only when they are presented with 'if… then…' sentences and when the subject cannot rely on any extra information, such as contextual information or knowledge about the topic.

1. Introduction

1.1. Wason's selection task and conditional reasoning

Since Aristotle, logic has been conceived as the study of sound reasoning and the instrument (*organon*) on which sciences are based. Logicians show how idealized rules of reasoning work, focusing on the form rather than the content of arguments. But what is known about human reasoning in actual practice? Logical rules do not always reflect how people reason in the various circumstances of everyday life: we often have to make decisions under uncertain conditions and do so by appealing to heuristics rather than logic (Kahneman & Tversky 1981); rigorous thinking is not always essential. However, a long philosophical tradition holds that men and women are essentially rational beings and depend on a certain acquaintance with the basic rules of logical systems. The question then is what empirical data is available to support our belief in rationality. Over the last century, many studies in the psychology of reasoning have suggested that such a belief is wrong, since even trivial logical reasoning tests prove to be difficult for the subjects involved. The most famous of these tests is Wason's Selection Task (Wason 1966).[1] In his original experiment, the British psychologist Peter Wason put four cards in front of his subjects:

E	K	4	7

Figure 1. Wason's four cards.

Each card had a visible face and a hidden one. Under these circumstances, subjects were asked to judge whether the following sentence was valid or not:

[1] See also Wason & Johnson-Laird (1972) for a detailed presentation of the task.

(1) If a card has a vowel on one face, then that card has an even number on the opposite face.

In order to solve the task, subjects were allowed to turn only those cards whose hidden side they considered relevant to the truth or falsity of the sentence. This task is seen as an example of conditional reasoning, because the sentence consists of a claim introduced by 'then' (*subsequent*) under a certain condition introduced by 'if' (*antecedent*). Wason offered a logical analysis of the problem, arguing that, if we take P for 'Card x has a vowel on one face' and Q for 'Card x has an even number on the opposite face', we are left with the formula $P \supset Q$, a material implication belonging to the language of propositional logic with the following truth conditions:

P	Q	$P \supset Q$
True	True	True
True	False	False
False	True	True
False	False	True

Table 1. Truth conditions for material implication.

Wason's task is therefore solved only if the cards E and 7 are turned over. These two cards are both necessary and sufficient to verify the validity of sentence (1). Card E corresponds to the case in which P is true and card 7 corresponds to the case in which Q is false, which leaves open these possibilities for their hidden faces:

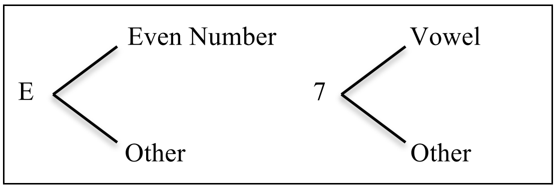

Figure 2. Open possibilities for cards E and 7.

If there is not an even number behind card E or there is a vowel behind card 7, the material implication $P \supset Q$ is invalidated, because we are in the situation defined by the second row in Table 1. It is easy to prove that the other two cards are not relevant because they leave open only possibilities incompatible with that row of the table, i.e., they can never falsify the rule.

In Wason's original experiment, fewer than 10% of subjects identified the correct solution, which suggested that humans are generally unable to follow basic logical rules in their reasoning. Moreover, Wason noted that most subjects chose card E (in some cases together with card 4), whilst very few people chose card 7. He explained this by making a distinction between confirmation biases and falsification biases, arguably inherited from Popper's (1959) philosophy of science: when deciding with the validity of a conditional sentence, people pay more attention to empirical data that confirm the sentence than to data that falsify it. The lack of a falsification bias in human cognition is, for Wason, the main flaw in our rationality.

1.2. Reasoning in context

Wason's selection task addresses an abstract problem through interactions of signs on cards, whose difficulty may be due to a lack of familiarity with the context. Similar considerations have led several psychologists to propose revised versions of the task in which subjects are expected to be more familiar. The main proposals in the literature are of two kinds: normative and ecological. Normative tasks involve permissions and obligations (i.e. deontic notions) rather than mere descriptions; Wason's selection task is not normative since it has nothing to do with deontic modalities. The difference can be explained by comparing (1) with the following examples of normative tasks (from Cheng & Holyoak 1985 and Manktelow & Over 1991):

(2) If an envelope is sealed, then it must have a 20 cent stamp.

(3) If you clear up spilt blood, you must wear rubber gloves.

The normative status of (2) and (3) derives from the fact that the modal verb (*must*) introduces a precondition to be satisfied before an action is taken. Clearly, this interpretation is not possible in the case of (1). However, it is important to note that the normative status is not merely due to the presence of a modal verb: it is the content of a proposition, together with the context in which it is uttered, that gives the modal a deontic sense.[2]

Ecological tasks are often used by researchers favouring an evolutionary approach in psychology, for example Cosmides (1989) and Cosmides and Tooby (1991). These authors hold that the ability to solve a given problem depends on whether the problem has been relevant for natural selection in the history of mankind. This means that some problems are adaptive and humans have developed *evolutionary algorithms* in order to solve them and survive, while other problems are non-adaptive and there is no specific cognitive ability to address them. Wason's selection task belongs to the latter group. Ecological alternatives to Wason's task usually address problems in a context of social exchange, i.e., where there are some benefits one receives after paying some costs. Here are some examples from Griggs and Cox (1982) and Cosmides (1989):

(4) If a person is drinking beer, then he must be at least 16 years old.

(5) If a student is to be assigned to Groover High School, then that student must live in Groover City.

(6) If a man eats cassava root, then he must have a tattoo on his face.

Most people know that a person can drink beer only under certain age conditions, so a social-contract interpretation for (4) is almost unavoidable. Cases (5) and (6) are more ambiguous, but if the experimental subjects are told that being assigned to Groover High School and eating cassava are beneficial, the social contract-interpretation applies again. According to Cosmides, in a context of social relations one has to beware of cheaters, i.e. people involved in a contract who try to reap benefits without paying any costs.

It has been shown that normative and ecological (i.e. concrete) tasks improve conditional reasoning, but the explanation of this phenomenon is controversial. For example, Cheng and Holyoak (1985) claim that in normative contexts people reason according to specific rules (*pragmatic schemas*) that are not available in descriptive ones. For Cosmides (1989), it is only the procedure detecting cheaters that improves conditional reasoning. This evolutionary thesis is supported by the fact that some deontic tasks unrelated to social contracts have a high rate of difficulty. However, Manktelow and Over (1991) show that in other cases the absence of a social-contract context is not problematic for experimental subjects. This suggests that a different kind of analysis is required.

1.3. Logic and natural language

The approach to conditional reasoning taken in this paper is built on the relationship between formal and natural language. Most difficulties with Wason's selection task, we argue, are not due to the absence of rigorous inferential mechanisms in the human cognition, but rather to

[2] An in-depth linguistic perspective on this topic may be found in Kratzer (1981). For a logical account of modalities, see Lewis & Langford (1932) and Hughes & Cresswell (1968).

difficulties in translating logical formulas into natural language. According to Wittgenstein (1922), the meaning of a logical operator is given by its truth conditions; there are logical operators, such as conjunction (&) and negation ($\sim$), whose translation into natural language gives a clear account of truth conditions, while other operators, such as material implication ($\supset$) and disjunction ($\vee$), are usually rendered in opaquely. The point is that the comprehension of a logical operator requires an exhaustive account of its truth conditions. Consider, for instance, the case of material implication: a natural language sentence of the form *If P then Q* does not actually reflect the truth conditions of the logical formula $P \supset Q$. In the first case, the only information is that under condition P also condition Q applies, while in the second case there is a symbol with broader implications. Accordingly, the truth table of material implications can be summarised in two sentences:

(7) If condition *P* applies, condition *Q* applies as well.

(8) If condition *Q* does not apply, condition *P* does not apply either.

Sentence (8) is often absent in natural language translations of $P \supset Q$, as in the case of Wason's original task; this justifies the fact that subjects do not thoroughly interpret the truth conditions of the intended formula but only perpetuate a confirmation bias, in line with their comprehension of clause (7). Moreover, the formula $P \supset Q$ is usually interpreted as having the truth conditions stipulated by the ancient philosopher Philo of Megara (those illustrated above), although another interpretation, by Diodorus Cronus, was also available at the time and has been revived in modal logic during the last century (cf. the notion of strict implication in Lewis & Langford 1932). Nowadays the logic of conditionals is a specific field of formal logic with many different approaches (see Jackson 1991).

This study aims therefore to account for factors influencing the solution of problems that involve natural language translations of logical formulas. Subjects can be expected to reason correctly through logical formulas only when the natural linguistic translation of such formulas is transparent with respect to truth conditions. This transparency is influenced by multiple factors: first of all, some operators have a counterpart in natural language, while others do not. Secondly, the information embedded in the scenario and the information in the subjects' memory (i.e. previous world knowledge) play an important role. Our hypothesis is that formulas can be translated when at least one of two conditions apply: the formula has a precise counterpart in natural language; there is some information the subject can rely when performing the translation. This implies that translations are always possible, unless the operator has no precise counterpart in natural language and there is no extra information for the subject to rely on.

2. Methodology

To test our assumption that truth condition transparency is of central importance in selection tasks, we conducted an experiment in Wason's style. The experiment was designed to show that a given problem may be solved more or less efficiently, depending on (i) the logical structure of the problem and (ii) the availability of extra information.

2.1. Participants

Thirty subjects were recruited by wall advertising in the Department of Communication at the University of Siena (mean age, 22.15 years; SD, 1.85; 18 females, 12 males). The study was conducted in Italian and all participants were native speakers of Italian, were students and came from different regions of Italy.

2.2. Stimulus material

Six different selection tasks were designed. Each had one of three possible implicit logical structures and consisted of an Italian sentence whose validity had to be checked in a certain context. The context was either abstract (finding interactions between signs on cards) or concrete (finding interactions between preconditions and benefits relating to the real world). The implicit logical structure of the problem embodied one of the following formulas: $P \supset Q$; $\sim(P \,\&\, \sim Q)$; $(\sim P \vee Q)$.

According to classical propositional logic, these formulas have the same truth condition: they are always true except when P is true and Q is false. Hence, the solutions of the six tasks were equivalent. The sentences used in the experiment are listed below, together with their English translation, context and implicit logical structure.

(9) Se una carta ha una vocale su un lato, allora ha un numero pari sull'altro.
[If a card has a vowel on one side, then that card has an even number on the opposite side]
Context = Abstract Structure = $P \supset Q$

(10) Se qualcuno beve birra, allora ha almeno 16 anni.
[If one drinks beer, they are at least 16]
Context = Concrete Structure = $P \supset Q$

(11) Non si dà il caso che ci sia una vocale su un lato e non ci sia un numero pari sull'altro.
[It is not possible for there to be a vowel on one side and not an even number on the other]
Context = Abstract Structure = $\sim(P \,\&\, \sim Q)$

(12) Non è permesso che qualcuno beva birra e non abbia almeno 16 anni.
[One is not allowed to drink beer if one is not at least 16 years old]
Context = Concrete Structure = $\sim(P \,\&\, \sim Q)$

(13) O non c'è una vocale su un lato, o almeno c'è un numero pari sull'altro.
[Either there is no vowel on one side, or at least there is an even number on the other]
Context = Abstract Structure = $(\sim P \vee Q)$

(14) O non si beve birra, oppure si ha almeno 16 anni.
[Either one does not drink beer, or one is at least 16 years old]
Context = Concrete Structure = $(\sim P \vee Q)$

Sentence (9) was taken from the original task in Wason (1966) and sentence (10) was taken from Cosmides (1989). The remaining four sentences were created by changing the structure of (9) and (10) in order to express the other logical formulas considered.

2.3. Procedure

Instructions were given orally in Italian and the task was then performed on a computer, where the subject could read the sentence and view the context. For the abstract version of the task, subjects were presented with Wason's four cards (Figure 1) and with one of the following sentences: 9, 11, 13. In the concrete version, pictures of four men were shown with a tag giving their age or drink type: *21*, *15*, *Coke*, *beer*. One of sentences 10, 12 or 14 appeared alongside the picture.

For the abstract version of the task, subjects were asked: "Which cards would you turn in order to understand whether the following statement is true? [followed sentence 9, 11 or 13]. Select only the cards you really need to solve the task".

For the concrete version, subjects were presented with the four pictures and introduced to the context by "You have to understand if the following statement is true [followed by sentence 10, 12 or 14]. You can ask me about the drinking habits of the young man or the old man, who ordered beer and who ordered Coke, but only ask me the questions you really need".

The answer was spoken aloud to the researcher, who typed it into a different computer. There was no time limit, and the subjects were allowed to give ‘No answer’ to their question. Correct answers were coded with 1, wrong answers with 0, no answer with 3. Each subject was presented with both the concrete and the abstract conditions of the same logical formula, as detailed in the table below.

Subjects	$P \supset Q$ *abstract*	$P \supset Q$ *concrete*	$\sim(P \,\&\, \sim Q)$ *abstract*	$\sim(P \,\&\, \sim Q)$ *concrete*	$(\sim P \vee Q)$ *abstract*	$(\sim P \vee Q)$ *concrete*
1-10	x	x	-	-	-	-
11-20	-	-	x	x	-	-
21-30	-	-	-	-	x	x

Table 2. Order of presentation.

Subjects 1-10 were presented with the abstract and concrete condition of $P \supset Q$; subjects 11-20 with the abstract and concrete condition of $\sim(P \,\&\, \sim Q)$; subjects 21-30 with the abstract and the concrete condition of $(\sim P \vee Q)$. The sequence was balanced so that five subjects in each group were presented with the abstract form followed by the concrete form, and five with the concrete followed by the abstract.

3. Results and discussion

The graph below illustrates the overall scores (number of correct answers per condition) in the experiment. Every condition was presented 10 times, so every coloumn represents the raw score out of 10. For instance, while the abstract condition of the first version was solved only once, the concrete condition of the same version was solved 8 times.

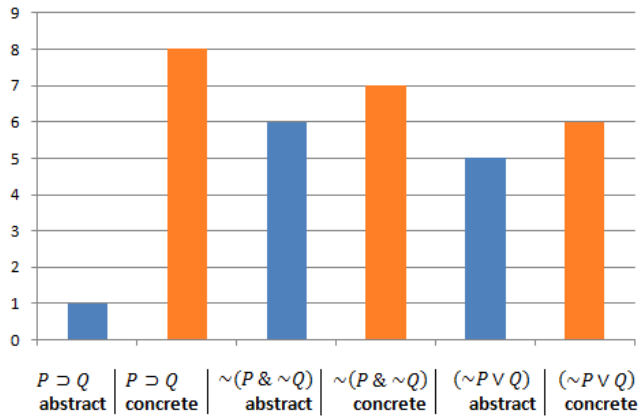

Figure 3. Raw scores per condition (out of ten).

At this point, three McNemar tests (one for each logical formula) were performed, comparing abstract vs. concrete conditions:

$P \supset Q$	χ^2 (1), p = .016
$\sim(P \,\&\, \sim Q)$	χ^2 (1), p = 1
$(\sim P \vee Q)$	χ^2 (1), p = 1

Then, two Cockran Q tests (one for abstract conditions alone and one for concrete conditions alone) were performed: χ^2 (2) = 10.33, p = .006; χ^2 (2) = 3.00, p = .22. Since there were three abstract conditions and three concrete conditions, six further Chi-squares were performed:

Abstract

$P \supset Q$ vs $\sim(P \,\&\, \sim Q)$ χ^2 (1) = 5.49, p = .019
$P \supset Q$ vs $(\sim P \vee Q)$ χ^2 (1) = 3.81, p = .051
$\sim(P \,\&\, \sim Q)$ vs $(\sim P \vee Q)$ χ^2 (1) = .202, p = .653

Concrete

$P \supset Q$ vs $\sim(P \,\&\, \sim Q)$ χ^2 (1) = .267, p = .606
$P \supset Q$ vs $(\sim P \vee Q)$ χ^2 (1) = .952, p = .329
$\sim(P \,\&\, \sim Q)$ vs $(\sim P \vee Q)$ χ^2 (1) = .202, p = .639

It is clear that the only significant differences are those including material implication in its abstract form. In all other cases, the difference is not significant (p > .05).

The task originally proposed by Wason (i.e. the abstract version of $P \supset Q$) proved significantly more difficult for subjects than any other version. However, this was not due to abstractness, as other abstract tasks with more transparent truth conditions were more readily solved. This suggests that the content of a sentence is not the main factor in evaluating performance, contrary to what is claimed in the theory of pragmatic schemas and in evolutionary psychology. We contend that the problem with Wason's selection task is that it does not reveal the truth conditions of the intended logical formula.

But what about the concrete counterpart of that task? How are we to explain the subjects' performance when assessing the validity of (10)? It seems that there are cases where a solution is reached even when the conditional statement presented is not an exhaustive translation of the truth conditions for material implication. As observed in Section 2, most of these cases have a deontic or social-contract context.

Let us now address the following issue: is it possible to give an explanation based on the transparency of truth conditions even for such tasks? We argue that performance in different contexts does not depend on content-specific algorithms, but rather on the kind of information available. When subjects try to understand the truth conditions of a task, they use all the information given in the scenario and stored in their memory, not only the content of the sentence that is assumed to be relevant. If part of the information about the truth table of an operator is not directly present in the sentence being tested and the task is easily solved, there are two possibilities: either the subject's memory holds the elements needed to integrate the meaning (as in the case of the interaction between beer and age) or this information is found in the text that accompanies the statement.

A similar case is found in Cosmides' (1989) two tests employing a fictitious cultural context to assess the validity of sentence (15):

(15) If you eat duiker meat, then you have found an ostrich eggshell.

Subjects were asked to think of themselves as anthropologists studying an exotic culture and collecting data from local boys on the relationship between eating duiker meat and finding ostrich eggshells. The first of Cosmides' two tests was conceived as a social-contract (i.e. ecological) task, the second as a descriptive task. A considerable difference was noticed in the subjects' performance, which was much higher in the first case. We argue that this may be due to additional information provided in the task under ecological conditions, whose text accompanying sentence (15) contained the following statements:

(16) Eating duiker meat is a privilege that must be earned.

(17) Finding ostrich eggshell is a sophisticated and difficult task.

(18) See if any of these boys have broken the law.

From this information, we infer that if a boy has not earned a certain privilege, he is not allowed to receive it. More specifically, it suggests the following idea:

(19) If one does not pay the cost of finding an ostrich eggshell, he does not receive the privilege of eating duiker meat.

Following Wason's procedure, we can take P for *X eats duiker meat* and Q for *X found an ostrich eggshell*, leading to the formula $\sim Q \supset \sim P$, with the following truth conditions:

P	Q	$\sim P$	$\sim Q$	$\sim Q \supset \sim P$
T	T	F	F	T
T	F	F	T	F
F	T	T	F	T
F	F	T	T	T

Table 3. Truth conditions for $\sim Q \supset \sim P$.

According Table 3, the subjects asked to check if any boy broke the law have to assess what happens if they did not find an ostrich eggshell ($\sim P$) or if they ate duiker meat (Q): they have to pay attention to the fact that the situation is not the one described in the second row of the table. Our previous comment on Wason's test also applies here: faced with an 'if…then…' sentence, subjects appear to think only of cases where the antecedent is true, while ignoring cases where the consequent is false. This is reasonable, if we accept that the 'if…then...' schema is not a comprehensive representation of material implication in natural language (as argued in Section 1.3). In the Cosmides social-contract task, however, subjects had the information of both (15) and (19): regarding (15), they chose to assess the situation where one eats duiker meat, i.e. P; while regarding (19) they chose to check the circumstances where one found no ostrich eggshell, i.e. $\sim Q$. Taking these choices together, the subjects solved the task because P and $\sim Q$ are the only situations able to falsify both $P \supset Q$ and $\sim Q \supset \sim P$. The two material implications here are logically equivalent, because they are true and false under the very same circumstances; however, when sentence (15) is presented together with other information leading to (19), the truth conditions for the logical structure $P \supset Q$ are fully provided, whereas when (15) is presented without such additional information, the truth conditions for $P \supset Q$ are only partially provided.

4. Conclusions

As outlined in Section 1.2, reasoning based on content-specific algorithms does not allow us to understand why in certain cases the difference between an abstract and a concrete version of a task is relevant, while in other cases it is negligible. Our experiment, for instance, shows that sentences with the logical structure $\sim (P \& \sim Q)$ or $(\sim P \vee Q)$ do not pose a serious problem for human reasoning either in abstract or in concrete contexts, while the situation is different in its material implications. We have attempted an analysis of a different kind, testing the assumption that reasoning abilities are mostly affected by the transparency of truth conditions. This seems to be confirmed by the fact that conditional problems are often presented in an opaque way (i.e. with an 'if… then…' sentence) and that subjects can be expected to perform well in such tasks only when further information on their truth conditions is available, as in certain deontic and ecological contexts.

What remains to be is explained how subjects actually process these tasks. Future research could address this question using online methodologies, such as eye tracking. It could be hypothesised that subjects that solve the task are not insensible to the confirmatory bias and (unlike those who do not solve the task) are not satisfied with the answer and go on to further processing. In other words, eye-tracking analysis might show a similar initial pattern for all subjects, whether they solve the task or not.

References

Cheng, P., & Holyoak, K. (1985). Pragmatic reasoning schemas. *Cognitive Psychology 17*, 391-416.

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

Cosmides, L., & Tooby, J. (1991). Reasoning and natural selection. In Dulbecco, R. (ed.) *Encyclopedia of Human Biology.* Vol. 6. San Diego: Academic Press, 493 -503.

Griggs, R., & Cox, J. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology 73*, 407-420.

Hughes, G., & Cresswell, M.J. (1968). *An Introduction to Modal Logic*. London: Methuen.

Jackson, F. (ed.) (1991). *Conditionals*. Oxford: Clarendon Press.

Kahneman, D., & Tversky, A. (1981). Judgment under uncertainty: heuristics and biases. *Science 185*, 1124-1131.

Kratzer, A. (1981) The notional category of modality. In Eikmeyer, H.-J., & Rieser, H. (eds) *Words, Worlds, and Contexts: New Approaches in Word Semantics*. Berlin: Walter de Gruyter, 38-74.

Lewis, C.I., & Langford, C.H. (1932). *Symbolic Logic*. New York: Dover Publications.

Manktelow K., & Over, D. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition 39*, 85-105.

Popper, K. (1959). *The Logic of Scientific Discovery*. London: Hutchinson. First published in 1935 as *Logik der Forschung* (Vienna: Springer Verlag).

Wason, P. (1966). Reasoning. In Foss, B.M. (ed.) *New Horizons in Psychology*. Harmondsworth: Penguin, 273-281.

Wason, P., & Johnson-Laird, P. (1972). *Psychology of Reasoning: Structure and Content.* Cambridge, MA: Harvard University Press.

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. London: Routledge & Kegan Paul. First published in 1921 as 'Logisch-Philosophische Abhandlung' in *Annalen der Naturphilosophische 14*.

Luca Cilibrasi is a PhD student in Clinical Language Sciences at the University of Reading. He holds a B.A. in Communication Studies and an M.A. in Linguistics and Cognitive Studies from the University of Siena. Email: l.cilibrasi@pgr.reading.ac.uk

Matteo Pascucci is a PhD student in Computer Science at the University of Verona. He holds a B.A. in Philosophy (University of Siena) and an M.A. in Philosophy and Logic (University of Trent). Email: matteo.pascucci@univr.it.