THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2009, 62 (3), 513–530

Psychology Press
Taylor & Francis Group

Response demands and the recruitment of heuristic strategies in syllogistic reasoning

Carlo Reverberi, Patrice Rusconi, Eraldo Paulesu, and Paolo Cherubini
University of Milan-Bicocca, Milan, Italy

Two experiments investigated whether dealing with a homogeneous subset of syllogisms with time-constrained responses encouraged participants to develop and use heuristics for abstract (Experiment 1) and thematic (Experiment 2) syllogisms. An atmosphere-based heuristic accounted for most responses with both abstract and thematic syllogisms. With thematic syllogisms, a weaker effect of a belief heuristic was also observed, mainly where the correct response was inconsistent with the atmosphere of the premises. Analytic processes appear to have played little role in the time-constrained condition, whereas their involvement increased in a self-paced, unconstrained condition. From a dual-process perspective, the results further specify how task demands affect the recruitment of heuristic and analytic systems of reasoning. Because the syllogisms and experimental procedure were the same as those used in a previous neuroimaging study by Goel, Buchel, Frith, and Dolan (2000), the result also deepen our understanding of the cognitive processes investigated by that study.

Keywords: Reasoning; Deduction; Syllogisms; Heuristics; Atmosphere effect; Belief bias; Mental models; Neuroimaging.

Categorical syllogisms are deductive arguments composed of three quantified propositions: two premises and a conclusion. The three propositions can be either particularly or universally quantified and can be either negative or affirmative, resulting in four types of syllogistic proposition: universal affirmatives (i.e., all A are B), universal negatives (i.e., no A are B), particular affirmatives (i.e., some A are B), and particular negatives (i.e., some A are not B). The two premises share a common term (“middle term”: e.g., in “all A are B”, “no B are C”, the middle term is B). The conclusion states a relationship between the two other terms (“extreme terms”). The syllogism is valid if and only if the conclusion necessarily follows from the premises. For more than a century, syllogisms have been widely used by empirical psychologists as an arena for the study of human deduction, and we now know many details regarding how people solve them.

People may sometimes recruit an array of analytic, logically consistent formal strategies,

Correspondence should be addressed to Paolo Cherubini, Department of Psychology, Università Milano-Bicocca, Piazza Ateneo Nuovo, 1, 20126, Milan, Italy. E-mail: paolo.cherubini@unimib.it

We thank Tom Ormerod and two anonymous referees for their helpful comments. For the same reason, and for having provided the original materials, we thank Vinod Goel.

http://www.psypress.com/qjep
DOI:10.1080/17470210801995010

grounded on the mental representation and processing of the logical structure of the premises (Bacon, Handley, & Newstead, 2003; Bucciarelli & Johnson-Laird, 1999; Ford, 1995; Roberts, 2000; Störring, 1908). Analytic strategies can be either semantic or syntactic—the former involving the construction of an exhaustive set of models representing the meaning of the premises (Erickson, 1974; Johnson-Laird, 2001; Johnson-Laird & Bara, 1984; Johnson-Laird & Byrne, 1991), the latter consisting in the application of truth-preserving transformations to the linguistic structure of the premises (Braine & O'Brien, 1998; Rips, 1994).

Yet, analytic strategies may be cognitively cumbersome (e.g., Gilhooly, Logie, & Wynn, 1999). In many instances, people trade logical accuracy for increased speed and decreased cognitive load. In these cases, they can recruit many different heuristic strategies: that is, fast and easy strategies, which offer no guarantee of a correct conclusion. Some heuristic strategies depend upon the activation of previous knowledge and can be applied only to syllogisms with thematic contents. The so-called belief bias is a tendency to eschew the requirement of ascertaining the logical necessity of a conclusion, by relying only on its believed truth or falsity. Other heuristic strategies do not depend on previous knowledge, but on an oversimplified, logically incomplete way of dealing with the structure of the premises. Examples abound: the atmosphere heuristic (Begg & Denny, 1969; Chapman & Chapman, 1959; Sells, 1936; Woodworth & Sells, 1935), the matching heuristic (Wetherick & Gilhooly, 1995), and building an incomplete set of models representing the premises (Johnson-Laird, 2001; Johnson-Laird & Bara, 1984; Ormerod, 2000; Ormerod & Richardson, 2003) can all be considered heuristic strategies loosely based on the structure of the premises. Contrary to knowledge-based heuristics, structural heuristics are available both for thematic and for abstract syllogisms. We describe one of them—the atmosphere heuristic—in the next paragraph. Overall, more than a hundred years of investigations portray a multistrategy view of syllogistic reasoning: People may engage either in analytic thinking (System 2, according to the dual-process perspective described by Stanovich & West, 2000) or in heuristic thinking (System 1). In the latter case, people sometimes rely on knowledge-driven, context-dependent heuristics, and at other times they use heuristics that deal with the structure of the premises. An aim of the present study is to clarify whether different task demands affect recourse to analytic or heuristic reasoning, and—for thematic problems—the relative contributions of structural and knowledge-driven heuristics.

Directly tied to the first goal, the second aim of this study is to estimate the extent to which analytic reasoning, structural heuristics, and knowledge-driven heuristics might have been involved in a series of studies on the neural basis of syllogistic reasoning by Vinod Goel and his colleagues (Goel, Buchel, Frith, & Dolan, 2000; Goel & Dolan, 2003, 2004). The authors found different neural systems involved in thematic and abstract syllogisms. The two systems share common components (left inferior prefrontal cortex, right cerebellum, and bilaterally fusiform gyrus and basal ganglia nuclei), but thematic reasoning engaged more the left superior temporal gyrus, whereas abstract reasoning engaged more the left superior parietal lobe. According to the authors, reasoning about syllogisms depicting familiar situations utilizes context-specific, knowledge-driven heuristics, whereas reasoning about abstract problems or unfamiliar situations utilizes formal methods. However, their studies imposed time constraints and used a peculiar set of syllogisms that could have encouraged recourse to heuristic reasoning even for abstract syllogisms. Thus, a careful disentangling of the relative roles of structural heuristics, knowledge-driven heuristics, and analytical reasoning would foster a better understanding of their findings.

Properties of the syllogisms investigated by Goel and colleagues (2000)

Goel et al. (2000) used 14 easy syllogistic premises, basing the evaluation of difficulty on Dickstein's

Table 1. *The structures of the arguments used by Goel et al. (2000)*

Premise 1	*Premise 2*	*Conclusion (valid)*	*Conclusion (invalid)*
All B are A	All C are B	All C are A (94.7)	No C are A (0)
All A are B	No C are B	No C are A (94.7)	Some C are A (0)
All A are B	No B are C	No C are A (89.5)	Some C are A (5.3)
All B are A	Some C are B	Some C are A (89.5)	No C are A (5.3)
All B are A	Some B are C	Some C are A (89.5)	No C are A (2.6)
All A are B	Some C are not B	Some C are not A (73.7)	(not used)
No B are A	All C are B	No C are A (92.1)	All C are A (0)
No A are B	All C are B	No C are A (92.1)	Some C are A (0)
No B are A	No B are C	(not used)	Some C are A (0)
No A are B	Some B are not C	(not used)	All A are C (0)
Some B are A	All B are C	Some C are A (84.2)	All C are A (2.6)
Some A are B	All B are C	Some C are A (89.5)	No C are A (2.6)
Some B are not A	All B are C	Some C are not A (78.9)	No C are A (2.6)
Some B are not A	All C are B	(not used)	No C are A (13.2)

Note: Average proportions of participants who selected that conclusion in Dickstein's (1978) study are reported in parentheses.

(1978) survey. Each couple of premises was matched with a valid conclusion, an invalid conclusion, or both, resulting in 24 syllogistic arguments; some of them were presented more than once, so that each participant was given 30 syllogisms (15 valid and 15 invalid) with abstract content and 30 structurally identical syllogisms with thematic content: 15 with a valid true conclusions or invalid false conclusions (congruent syllogisms); 15 with valid false conclusions or invalid true conclusions (incongruent syllogisms). Their structures are reported in Table 1 (V. Goel, personal communication, October 12, 2004).

The first premise was displayed for 3 s before adding the second premise. The conclusion was shown 3 s after the second premise had been displayed. Participants were allowed a randomly selected time interval (ranging from 3.75 s to 7.85 s) for deciding whether the conclusion followed necessarily from the premises. At the start of each trial, the participants did not know whether they were allowed 3.75 s or more to respond; furthermore, they were instructed to respond as fast as possible. Accordingly, they acted under strict time constraints. Henceforth, participants probably sought simple strategies that could generate viable responses as fast as possible. Notably, all the syllogisms (Table 1), both the abstract and the thematic syllogisms, could be solved correctly by a very easy, nonlogical, heuristic based on the atmosphere effect (Begg & Denny, 1969; Chapman & Chapman, 1959; Sells, 1936; Woodworth & Sells, 1935).[1] The heuristic principles that generate this effect are as follows:

1. If at least one of the premises is negative, then the conclusion is negative; otherwise, it is affirmative.
2. If at least one of the premises is particular, then the conclusion is particular; otherwise, it is universal.

These rules qualify as heuristic thinking according to the description made by most dual-process views of reasoning (e.g., Evans, 1984;

[1] The 15 syllogisms used in the 2003 study by Goel and Dolan differed from those used in 2000 by Goel et al. for four items. However, even for those syllogisms the correct solution coincides with the solution suggested by an atmosphere-based heuristic strategy. Of the 24 syllogisms used in the 2004 study, 15 had a valid conclusion consistent with atmosphere, 1 had a valid conclusion not consistent with atmosphere (ae1), and the rest were invalid. Whether invalid syllogisms are consistent or not with atmosphere depends on the specific conclusion proposed to the participants. This information is not available in the paper. Both in the 2003 and in the 2004 studies the time constraint remained as in the 2000 study.

Evans & Over, 1996; Kahneman, 2003; Sloman, 1996; Stanovich & West, 2000; Wason & Evans, 1975): fast, intuitive, and effortless (De Neys, 2006a, 2006b) thinking processes that anchor their conclusions to one or a few superficial, easily accessible features of the premises. Even though they are not logically correct inference rules, atmosphere heuristics sometimes generate correct conclusions. For example, when applying the two principles to the set of premises:

All A are B
All B are C

they generate a universal-affirmative atmosphere for the conclusion that is correct. Nonetheless, at other times the atmosphere principles generate incorrect conclusions. For example, when applying the two principles to the following set of premises:

All B are A
All B are C

they generate the same universal-affirmative atmosphere as before, which is now incorrect because the valid conclusion is "some C are A" (or "some A are C"). The atmosphere principles fare reasonably well when dealing with valid syllogisms: Only five out of 27 couples of valid syllogistic premises do not have any conclusion consistent with their atmosphere.[2] However, there are 37 couples of premises that do not admit any valid conclusion. For these syllogisms the conclusions suggested by the atmosphere are always wrong. Overall, the atmosphere principles support incorrect conclusions for 42 couples of premises out of 64. Gilhooly and colleagues (Gilhooly, Logie, Wetherick, & Wynn, 1993; Gilhooly, Logie, & Wynn, 1999; Wetherick & Gilhooly, 1995) have estimated that up to 77% of participants in experiments using categorical syllogisms spontaneously adopt the atmosphere heuristics, or some similar logically incorrect principle (e.g., the "matching principle", which can also correctly solve all or most of the syllogisms used by Goel and colleagues).

Because of the success of this heuristic on most of the syllogisms presented by Goel et al. (2000; Goel & Dolan, 2003, 2004), we do not know whether the above-chance performance scores reported in those studies are mainly caused by recourse to logically consistent strategies or to structural heuristics (or both). It is plausible that under time pressure, many participants could have sought and applied fast heuristic strategies. Even those participants who were willing to engage in analytic processing at an initial stage could have abandoned it soon, switching to the easily available—and successful—heuristic shortcut, in keeping with previous literature showing that successful heuristic rules can be learned by exposition to sets of homogeneous problems and then mechanically applied to superficially similar new problems (Cherubini & Mazzocco, 2004; Luchins, 1942; Woltz, Gardner, & Bell, 2000). Thus, in this study we checked whether strict time constraints and exposition to an initial set of stimuli homogeneously consistent with the atmosphere heuristic induced widespread use of that heuristic (or equivalent heuristics) and whether relaxing those constraints increased the recourse to analytic reasoning strategies. Furthermore, by using abstract syllogisms (Experiment 1) and

[2] The five couple of premises with no conclusions consistent with the atmosphere are:

1. No A are B, all B are C (conclusion: Some C are not A).
2. No B are A, all B are C (conclusion: Some C are not A).
3. All B are A, no B are C (conclusion: Some A are not C).
4. All B are A, no C are B (conclusion: Some A are not C).
5. All B are A, all B are C (conclusion: Some A are C).

In this paper, like in most psychological studies on syllogistic reasoning and in many logical systems, we assume that the universally quantified premises have an existential entailment. In the logical systems where this assumption is rejected none of the syllogisms above entails any valid conclusion— for example, "all B are A, all B are C" does not entail "some A are C", because the latter is not necessarily true when the B set is empty.

thematic syllogisms (Experiment 2) we measured the relative roles of structural heuristics, knowledge-driven heuristics, and analytical reasoning, both in the time-constrained conditions and in the relaxed, unconstrained conditions.

EXPERIMENT 1: ABSTRACT SYLLOGISMS

Method

Design and predictions

We grouped all the possible syllogistic forms into four theoretical categories according to their logical validity and to the consistency of the conclusion with the atmosphere of the premises, as reported in Table 2.

Participants in the time-constrained condition received an initial set of 46 syllogisms (all of them belonging to Type 1 or 4), identical to the set used in Goel et al.'s (2000) study, with similar instructions and identical time constraints.[3] We describe these syllogisms as a training set, because, according to our conjecture, many participants exposed to this set should develop and then mechanically apply atmosphere-based or equivalent heuristics.[4] Immediately after receiving the training set, participants in the constrained condition received a test set comprising 62 syllogisms belonging to all four possible types: If the training test induced a widespread use of an atmosphere-based heuristic, then participants should answer correctly to Type 1 and Type 4 syllogisms in the test set (where the response suggested by the atmosphere is correct), but they should systematically fail with Type 2 and Type 3 syllogisms (where the response suggested by the atmosphere is incorrect): In other words, responses to Type 2 and Type 3 syllogisms in the test set of the constrained condition are an indicator of the use of analytic strategies (shown by correct responses) or heuristic strategies (incorrect responses). Participants in the unconstrained condition received 74 syllogisms of the four types, including all the syllogisms in Goel et al.'s stimuli set (2000), presented in random order. Participants in this condition were given additional support to engage in analytic reasoning: They were given a more extensive preexperimental practice period, were allowed to use paper and pencil, and could evaluate the arguments at their own pace. The unconstrained condition maximized resources available to tackle a heterogeneous set of syllogisms—where heuristics were not invariably correct. It tested whether participants performing under such conditions would be more likely to recruit fast heuristics or more effortful, time-consuming analytic reasoning processes with Types 2 and 3 syllogisms.

Table 2. *An example for each of the four types of syllogisms obtained by crossing validity of the conclusion with consistency of the conclusion with the atmosphere of the premises*

Conclusion is valid?	*Conclusion is consistent with atmosphere?* Yes	No
Yes	1. All B are A Some B are C Therefore, some A are C	2. All B are A All B are C Therefore, some A are C
No	3. All A are B Some B are C Therefore, some A are C	4. All B are A All C are A Therefore, some A are C

Note: All syllogisms used by Goel et al. (2000) and reported in Table 1 were of Type 1 or Type 4.

[3] We used 46 arguments, instead of 30 as in Goel et al. (2000), in order to have more data points to observe whether some learning occurred during the task (see the Results section).

[4] "Training set" is an appropriate name from the perspective of the experimenters. However, the training and test sets did not differ in any way from the perspective of the participants: There were no interruptions between the two sets, the stimuli presentation and response constraints were the same, there were no feedbacks, and so on. That is, the training set in the constrained condition should not be confounded with the explicit, preexperimental training with accuracy feedback administered to participants in the unconstrained condition.

If our conjecture is correct, participants in the unconstrained condition should not rely on an atmosphere heuristic as much as do participants in the constrained condition, as shown by an increased performance in Type 2 and Type 3 syllogisms in the unconstrained condition. Furthermore, under the assumption that more sophisticated reasoning strategies require more time, and that participants in the constrained condition were encouraged to adopt the atmosphere heuristic because of the severe time constraints imposed upon their responses, response latencies should be longer in the unconstrained than in the constrained condition. Specifically, in the constrained condition response latencies to arguments consistent and inconsistent with the atmosphere should be the same, whereas in the unconstrained condition, arguments inconsistent with the atmosphere (Types 2 and 3) should have longer latencies.

Participants
A total of 61 undergraduate students (46 female, 15 male; mean age 21.4 years, range 19–29 years; mean education 14.7 years, *SD* 2.1) of the University of Milan-Bicocca volunteered to participate to the study. None of them had taken any course in logic or the psychology of reasoning. A total of 35 participants were randomly assigned to the time-constrained condition and the remaining 26 to the unconstrained condition.

Procedure
The experiment was carried out in Italian under the control of a personal computer running the Presentation™ experimental software. Responses were given by pressing two keys on the computer keyboard, labelled "YES" and "NO". Participants were required to determine as fast as possible whether each conclusion was necessarily true according to its premises. The procedure in the constrained condition mimics that of Goel et al. (2000; V. Goel, personal communication, October 12, 2004). After the instructions, two arguments were shown and explained. Two practice trials were administered followed by feedback and were later excluded from the analyses. The beginning of a trial was signalled by an asterisk (*). The sentences appeared on the screen one at a time with the first sentence appearing at 500 ms, the second at 3,500 ms, and the last sentence at 6,500 ms. All sentences remained on the screen until the end of the trial. The length of trials varied from 10.25 to 14.35 s, leaving participants 3.75 to 7.85 s after the presentation of the third sentence to respond. The order of presentation of the arguments was randomized. The first 46 syllogisms, 22 Type 1 and 24 Type 4, were the training set (Table 1; most were presented twice). The remaining 62 syllogisms were the test set, comprising 15 Type 1 syllogisms, 15 Type 4 syllogisms, 12 Type 2 syllogisms, and 20 Type 3 syllogisms (5 of which admitted a valid conclusion different from the one displayed; the remaining 15 did not admit any valid conclusion). Time constraints during the test set were the same as those in the training set. The letter triples used as terms in the stimuli were randomly chosen.

Participants in the unconstrained condition made self-paced responses. After reading the instructions and two commented examples, participants received 28 practice trials with accuracy feedback, which included examples of all four types of syllogisms. The experimental trials were 74 arguments of all four types. Participants were explicitly told that they could use paper and pencil in order to work out their responses if they so wished.

Results

Table 3 shows the mean proportion of correct responses for the four types of syllogism.

The proportions of correct responses in the test set of the constrained and unconstrained conditions were analysed by a 2 × 2 × 2 mixed-design analysis of variance (ANOVA), within-participants factors being the atmosphere, in terms of consistency with the correct response (consistent, Type 1 and Type 4, vs. inconsistent, Type 2 and Type 3) and validity of the argument (valid, Type 1 and Type 2, vs. invalid, Type 3 and Type 4), and the between-groups factor being the condition (constrained vs. unconstrained).

Table 3. *Mean proportion of correct responses and mean response latencies for the four sorts of argument in the two conditions*

		Congruent response		*Incongruent response*	
		Valid (Type 1)	*Invalid (Type 4)*	*Valid (Type 2)*	*Invalid (Type 3)*
Correctness	Constrained condition, training set	.79 (.12)	.76 (.12)	(absent)	(absent)
	Constrained condition, test set	.85 (.10)	.84 (.16)	.46 (.31)	.12 (.10)
	Unconstrained condition	.86 (.10)	.90 (.09)	.75 (.23)	.43 (.21)
Latencies	Constrained condition, training set	2.1 (0.5)	2.1 (0.5)	(absent)	(absent)
	Constrained condition, test set	1.8 (0.6)	2.0 (0.5)	2.0 (0.6)	1.9 (0.5)
	Unconstrained condition	12.4 (6.1)	13.2 (7.0)	13.1 (6.3)	14.9 (7.3)

Note: Congruent/incongruent denotes whether the atmosphere-suggested response is congruent with the correct response. Valid/invalid denotes type of argument. Mean response latencies in s. Standard deviations in parentheses.

All the main effects were reliable—atmosphere, $F(1, 60) = 202.7$, $MSE = 0.052$, $p < .00001$, means = .86 (consistent), .44 (inconsistent); validity, $F(1, 60) = 67.1$, $MSE = 0.023$, $p < .00001$, means = .73 (valid), .57 (invalid); condition, $F(1, 60) = 48.1$, $MSE = 0.035$, $p < .00001$, means = .57 (constrained condition), .74 (unconstrained condition)—showing an overall tendency to respond correctly more often to items where the atmosphere suggested the right response, to valid arguments, and in the unconstrained condition.

Atmosphere interacted reliably with condition—$F(1, 60) = 20.4$, $p < .0001$, $MSE = 0.052$, means = .84 (consistent/constrained), .88 (consistent/unconstrained), .29 (inconsistent/constrained), .59 (inconsistent/unconstrained)—showing that performance in problems where the atmosphere suggested the wrong response is better in the unconstrained than in the constrained condition, whereas there is no difference across conditions where atmosphere suggests the right conclusion. There was also a reliable interaction between atmosphere and validity—$F(1, 60) = 78.6$, $MSE = 0.022$, $p < .00001$, means = .86 (consistent/valid, Type 1), .87 (consistent/invalid, Type 4), .61 (inconsistent/valid, Type 2), .28 (inconsistent/invalid, Type 3)—showing that participants were more willing to accept an invalid conclusion consistent with the atmosphere than to incorrectly reject a valid conclusion inconsistent with the atmosphere. Closer examination shows that, in both conditions, there were reliably more correct responses to Type 2 valid than to Type 3 invalid syllogisms—constrained, $t(35) = 6.8$, $p < .00001$; unconstrained, $t(25) = 7.0$, $p < .00001$—whereas Types 1 and 4 syllogisms were not reliably different. However, performances for both sorts of syllogism improved significantly in the unconstrained condition: Type 2, $t(25) = 5.5$, $p < .0001$; Type 3, $t(35) = -23.8$, $p < .00001$.

Individual performances were analysed to examine how biases towards correct or incorrect responses were distributed across the sample (Table 4).

The performance of each individual participant was compared to the chance level by a binomial two-tailed test. If the null hypothesis could be rejected ($p < .05$), the participant was classified as showing a reliable bias either toward the correct or toward the incorrect response. In order to perform the binomial tests on a reasonable number of data points, and since the relevant group averages and standard deviations were similar, we analysed all the problems where the atmosphere suggested the correct response together (Type 1 and Type 4). Type 2 and Type 3 problems could not be analysed together, because the group performance in them was not homogeneous. Since there were only 12 and 9 Type 2 problems in the constrained condition (test set) and unconstrained condition, respectively, we excluded Type 2 problems from the analysis. Table 4 shows that the proportion of participants with a reliable bias changes from .97 in the constrained condition to .28 in the

Table 4. *Proportion of participants showing a reliable bias towards correct and incorrect responses in each condition*

	Types 1 & 4				*Type 3*			
	N	*Bias toward correct*	*Bias toward incorrect*	*No bias*	*N*	*Bias toward correct*	*Bias toward incorrect*	*No bias*
Constrained condition, training set	46	.91	0	.09				
Constrained condition, test set	30	.94	0	.06	20	0	.97	0.03
Unconstrained condition	45	.96	0	.04	20	0.12	.28	0.64

Note: "*N*" refers to the number of syllogisms available for each statistic.

unconstrained condition. These findings, corroborating those at the group level, show that there is an almost universal bias towards incorrect responses on Type 3 syllogisms in the test set of the constrained condition.

Response latencies

Table 3 also shows latencies of responses to the four sorts of syllogism. Latencies were analysed by a 2 × 2 × 2 mixed-design ANOVA, within-participants factors being validity and atmosphere, and the between-groups factor being condition. All the factors had reliable main effects—atmosphere, $F(1, 60) = 6.7$, $MSE = 3.81\text{E}{+}6$, $p < .05$, means = 7.3 s (consistent), 8.0 s (inconsistent); validity, $F(1, 60) = 5.4$, $MSE = 4.80\text{E}{+}6$, $p < .05$, means = 7.3 s (valid), 8.0 s (invalid); condition, $F(1, 60) = 127.8$, $MSE = 1.56\text{E}{+}7$, $p < .00001$, means = 1.9 s (constrained), 13.4 s (unconstrained)—showing that responses were slower when the atmosphere did not suggest the right conclusion, when the argument was invalid, and in the unconstrained condition. The latter predicted effect shows that when people can comfortably reason at their own pace, with more chances of recruiting analytic strategies, then they spontaneously use more time, even for the simplest syllogisms (Type 1), than they use in the time-constrained condition.

Condition interacted reliably with atmosphere—$F(1, 60) = 5.8$, $MSE = 3.81\text{E}{+}6$, $p < .05$, means = 1.9 s (consistent atmosphere, constrained), 12.8 s (consistent atmosphere, unconstrained), 1.9 s (inconsistent atmosphere, constrained), 14.0 s (inconsistent atmosphere, unconstrained)—and with validity—$F(1, 60) = 5.0$, $MSE = 4.80\text{E} + 6$, $p < .05$, means = 1.9 s (valid, constrained), 12.7 s (valid, unconstrained), 1.9 s (invalid, constrained), 14.0 s (invalid, unconstrained). The former interaction shows that responses to atmosphere-inconsistent problems were slower than responses to atmosphere-consistent problems only in the unconstrained condition, consistent with the idea that in the unconstrained condition participants were more able to detect the problems where the atmosphere was misleading.

The syllogisms in the training set of the constrained condition were solved more slowly than those in the test set (2.1 s vs. 1.9 s), $t(35) = 2.95$, $p < .01$, further suggesting that participants in the test set were actually learning a strategy.

Discussion

As shown by the critical Types 2 and 3 syllogisms in the test set of the constrained condition, the atmosphere of the premises accounts for a majority of responses. Analyses of individual performances for Type 3 syllogisms show that almost all participants (97%) in the constrained condition were reliably biased toward the incorrect response suggested by the atmosphere. These results corroborate the idea that the prevailing form of reasoning in the constrained condition was heuristic reasoning, based on atmosphere or equivalent heuristics. Recourse to analytic strategies increased when more time and resources were available, shown by an increase in accuracy in the unconstrained condition for arguments whose correct response was inconsistent with the atmosphere—a trend present in the overall group performance (an average + 30% correct responses for Type 2 and Type 3 syllogisms) as well as in the distribution of individual performances.

In our view, two factors boosted the use of heuristic strategies in the constrained condition:

1. Responses were severely time constrained, encouraging the development and use of fast heuristic shortcuts. The spontaneous response time required for evaluating arguments was far longer than the time available in the constrained condition, as shown by the comparison of latencies in unconstrained and constrained conditions.
2. Participants faced an initial set of arguments whose correct responses were homogenously consistent with responses suggested by the atmosphere heuristic: In these circumstances, even those participants willing and able to cope analytically with the arguments would soon learn the atmosphere-based shortcut and then apply it mechanically to the following problems (Cherubini & Mazzocco, 2004; Luchins, 1942; Woltz et al., 2000).

In both conditions, people preferred accepting valid arguments rather than rejecting invalid arguments. This tendency interacted reliably with the atmosphere of the premises: It was exclusively present when atmosphere was not consistent with the correct response—that is, for Type 2 and Type 3 syllogisms. At least three interpretations are possible. First, a close scrutiny of the stimuli (see Footnote 3) shows that the proposed conclusions of Type 2 syllogisms were always the subordinate of the conclusions suggested by the atmosphere.[5] In other words, the invalid conclusions suggested by the atmosphere heuristic were either "all A are B" or "no A is B", whilst the proposed valid conclusions were "some A are B" or "some A are not B", respectively. In these conditions, the particular conclusion can be accepted either because—for heuristic reasons—it is more "cautious" than the corresponding universal conclusion (Begg & Denny, 1969; Chapman & Chapman, 1959), or—for logical reasons—because it follows straightforwardly from its universal counterpart (some statements to this effect were found in the notes written by the participants in the unconstrained condition).

The second interpretation capitalizes upon an analogous interaction, typically observed for thematic syllogisms: In some studies, previous beliefs affect invalid syllogisms more than valid syllogisms (Evans, Barston, & Pollard, 1983; Morley, Evans, & Handley, 2004). This trend suggests that some heuristic effects are stronger for invalid arguments: They raise the acceptance level of invalid conclusions more than they suppress the acceptance level of valid conclusions. In the present case, the correctness scores are compatible with the idea that a fast atmosphere-based process suggests either a "yes" or a "no" response. If a "yes" response is suggested (Types 1 and 3 syllogisms), it is most often accepted and produced without further control—causing the many correct Type 1 "yes" responses and almost as many incorrect Type 3 "yes" responses. If a "no" response is suggested (Types 2 and 4 syllogisms), before being produced it has a chance of being further checked by analytic processes: In the case of Type 4 syllogisms, analytic processes confirm the correct "no" response; however, in the case of Type 2 syllogisms they do not confirm it, thus lowering the proportion of incorrect "no" responses to Type 2 syllogisms.

The third interpretation capitalizes upon a proposal originally related to mental models theory: When dealing with multimodel syllogisms such as Types 2 and 3 in evaluation tasks, people may eschew building all the models of the premises and instead build a model of the conclusion, check whether the premises are compatible with it (i.e., map the premises onto the model of the conclusion), and—if so—accept the conclusion

[5] This design imbalance is due to a structural feature of categorical syllogisms themselves, rather than to a bias in our selection of the stimuli.

(e.g., Cherubini, Garnham, Oakhill, & Morley, 1998; Klauer, Musch, & Naumer, 2000; Morley et al., 2004). Formally, this backward strategy, observed in some of the diagrams depicted by participants in the unconstrained condition, amounts to reasoning about the possibility of the conclusion, disregarding its necessity (see also the "misinterpreted necessity" proposal by Evans et al., 1983). All conclusions to Types 2 and 3 syllogisms were possible (even though they were necessary—i.e., valid—only for Type 2 syllogisms). Accordingly, people adopting the backward strategy would respond correctly (for the wrong reasons) to Type 2 syllogisms, but incorrectly to Type 3 syllogisms.

The second interpretation implies that some analytic processing was involved in the constrained condition; the first and third interpretations account for the finding in terms of structural heuristics, either atmosphere (plus the logically correct intuition that universal conclusions entail their subordinate particular conclusions), or the building of oversimplified models of problems. Of course, it is entirely possible that the three interpretations refer to three alternative strategies and that all of them could be adopted by different participants in different trials.

Even allowing for the unpredicted effect of validity, in the test set of the constrained condition analytic strategies were scarcely recruited (29% correct responses, on average, to Types 2 and 3 syllogisms). In the unconstrained condition, recourse to analytic strategies was higher (58% correct responses, on average, to Types 2 and 3 syllogisms)—but it did not completely rule out heuristic strategies (42% of responses). The exploration of the notes jotted down by participants in this condition hinted at a variety of strategies. Of the 15 participants that used paper and pencil, 12 drew (at least for one syllogism) either diagrams similar to Euler circles, or connecting terms with "=" and "≠" (or similar symbols), suggesting the use of some kind of semantic, set-based strategy; 9 rewrote (at least for one syllogism) the premises and/or conclusion, commented on them, linked them, used arrows to represent exchange of terms or matching of quantifiers, or used similar representational formats, suggestive of the use of syntactic, linguistic strategies. In the cases where a conclusion was clearly stated in the notes, there are instances both of atmosphere-consistent errors and of correct answers associated to both classes of strategies.

EXPERIMENT 2: THEMATIC SYLLOGISMS

Experiment 1 showed that the dominant processes for solving a homogenous set of abstract Type 1 and Type 4 syllogisms under strict time constraints are structural heuristics, like atmosphere-based or equivalent strategies. Thematic syllogisms can also recruit different, knowledge-driven heuristic strategies, most notably belief bias (Morgan & Morton, 1944; Wilkins, 1928). Belief bias is a confounding between the truth and the validity of a conclusion: People prefer endorsing true rather than false conclusions, accepting more true valid than false valid conclusions, and rejecting more false invalid than true invalid conclusions (for details and theories, see Cherubini et al., 1998; Evans et al., 1983; Klauer et al., 2000; Morley et al., 2004; Newstead, Pollard, Evans, & Allen, 1992; Oakhill & Garnham, 1993; Oakhill & Johnson-Laird, 1985; Revlin, Leirer, Yopp, & Yopp, 1980). In the dual-process view, semantic processes are automatically activated each time the context associatively evokes prior knowledge (Sloman, 1996). Reliance on analytical reasoning requires a voluntary inhibitory effort aimed at suppressing beliefs that conflict with the validity of the conclusion. The inhibition of incongruent beliefs sometimes fails; in those cases, participants report the believed truth (or falsity) of the conclusion, instead of its validity. However, thematic syllogisms are also amenable to structural heuristics and analytic strategies, exactly as abstract syllogisms. We here investigate the relative contribution of analytic strategies, structural heuristics, and knowledge-driven heuristics in solving thematic syllogisms, both under strict time constraints (constrained condition) and in an unconstrained setting (unconstrained condition).

If the atmosphere heuristic plays a major role in the time-constrained condition, then atmosphere, instead of validity, should conflict with beliefs in that condition: Where both heuristic cues suggest the same conclusion, correct or incorrect, a majority of participants should accept it, and performance should be at its highest or lowest, respectively. Where one heuristic cue suggests one response, and the other cue suggests a different response, performance should tell us which heuristic process is the leading one. The same trend should be present in the unconstrained condition, but, since use of analytic strategies is more likely in that condition, a more typical conflict between beliefs and validity should surface. That is, the effect of beliefs should be modulated by validity in the unconstrained condition more than in the constrained one.

Method

Participants

A total of 45 undergraduate students (34 female, 11 male; mean age 22.5 years, range 20–43 years; mean education 15.0 years, *SD* 2.3) of the University of Milan-Bicocca participated to the study in exchange for course credits. A total of 23 participants were randomly assigned to the constrained condition and the remaining 22 to the unconstrained condition.

Design and procedure

The design and procedure were identical to those of Experiment 1, except for what follows. The Validity × Atmosphere design of Experiment 1 was expanded by adding a third orthogonal factor, belief congruency. We used the same thematic syllogisms as those of Goel et al. (2000); those with true-valid and false-invalid conclusions were "congruent" syllogisms, whereas those with true-invalid or false-valid conclusions were "incongruent" syllogisms. There were 80 syllogisms in the training set of the constrained condition: 40 Type 1 and 40 Type 4, half of them congruent and half incongruent. So, in the training set the atmosphere heuristic always suggested the correct answer (as in Experiment 1), but the belief heuristic suggested the correct answer only half the time. In the test set, there were 124 syllogisms: 30 Type 1 syllogisms, 30 Type 4 syllogisms, 24 Type 2 syllogisms, 40 Type 3 syllogisms (10 of which admitted a valid conclusion different from the one displayed and the remaining 30 that did not admit any valid conclusion). Half of each sort were belief congruent, the other half incongruent. The 154 syllogisms in the unconstrained condition comprised 40 Type 1, 50 Type 4, 24 Type 2, and 40 Type 3 (10 of which admitted a valid conclusion different from the one displayed and 30 of which did not admit any valid conclusion). Half were congruent, and half were incongruent. In both conditions, on completion of the syllogistic task, each participant was asked to evaluate each presented conclusion as true or false by pressing one of two labelled keys on the computer's keyboard. This manipulation check tested whether the participant's beliefs were consistent with the constrained categorization of belief-congruent and belief-incongruent syllogisms. A threshold of 70% accuracy was set for accepting a participant. In the debriefing session all participants were administered a questionnaire on reasoning strategies used by Bacon et al. (2003).

Results

Data from 2 participants (both from the unconstrained condition) were discarded before the analyses, because their evaluation of the truth of the conclusions was not consistent with our own. Table 5 shows the mean proportion of correct responses for the remaining 43 participants.

The proportions of correct responses were analysed by a 2 × 2 × 2 × 2 mixed-design ANOVA, within-participants factors being the atmosphere consistency, validity, and the congruence of prior beliefs. The between-groups factor was the condition. All the main effects were reliable—atmosphere, $F(1, 41) = 197.7$, $MSE = 0.08$, $p < .00001$, effect size (measured as partial eta squared) = .83, means = .88 (consistent), .44 (inconsistent); validity, $F(1, 41) = 44.1$, $MSE = 0.05$, $p < .00001$, effect size = .52, means = .74 (valid), .58 (invalid); beliefs, $F(1, 41) = 23.4$,

Table 5. *Mean proportion of correct responses for the eight sorts of thematic syllogism in the two conditions of Experiment 2*

		Correct		*Incorrect*	
Belief		*Valid (Type 1)*	*Invalid (Type 4)*	*Valid (Type 2)*	*Invalid (Type 3)*
Congruent	Constrained task, training set	.87 (.11)	.78 (.17)	(absent)	(absent)
	Constrained task, test set	.92 (.08)	.81 (.16)	.62 (.20)	0.25 (.15)
	Unconstrained condition	.94 (.07)	.93 (.07)	.65 (.29)	0.47 (.22)
Incongruent	Constrained task, training set	.79 (.10)	.65 (.21)	(absent)	(absent)
	Constrained task, test set	83 (.16)	.80 (.16)	.46 (.22)	.18 (.14)
	Unconstrained condition	.94 (.06)	.86 (.12)	.53 (.36)	.38 (.26)

Note: Correct/incorrect refers to the response suggested by the atmosphere. Valid/invalid denotes type of argument. Standard deviations in parentheses.

MSE = 0.02, $p < .0001$, effect size = .36, means = .70 (congruent), .62 (incongruent); condition, $F(1, 41) = 11.7$, *MSE* = 0.08, $p < .005$, effect size = .22, means = .61 (constrained), .71 (unconstrained)—showing that there was an overall tendency to respond more correctly to items where the atmosphere suggested the right response, to valid arguments, to items where previous beliefs were congruent with the response, and in the unconstrained condition (in order of effect sizes).

The predicted conflict between atmosphere and beliefs occurred, with the former dominating the latter, as shown by the Atmosphere × Beliefs interaction, $F(1, 41) = 16.5$, *MSE* = 0.007, $p < .0005$, means = .90 (consistent atmosphere/congruent beliefs), .86 (consistent atmosphere/incongruent beliefs), .50 (inconsistent atmosphere/congruent beliefs), .39 (inconsistent atmosphere/incongruent beliefs). We further analysed this interaction by rescoring the data, using as the dependent variable the proportion of acceptances of the conclusion (instead of the proportion of correct responses), and reshaped the two factors in terms of the response that they suggested: Atmosphere (suggesting "yes", Types 1 and 3] vs. "no", Types 2 and 4) × Beliefs (true conclusions suggest "yes", false conclusions suggest "no"). The interaction remained significant, $F(1, 41) = 9.9$, *MSE* = 0.013, $p < .005$, means = .81 (atmosphere "yes"/beliefs "yes"), .80 (atmosphere "yes"/beliefs "no"), .40 (atmosphere "no"/beliefs "yes"), .31 (atmosphere "no"/beliefs "no"), meaning that beliefs played a negligible role when the atmosphere suggested to accept the conclusion, independently from its correctness.

Returning to the standard correctness analyses, the three-way interaction between beliefs, validity and condition was significant, $F(1, 41) = 7.0$, *MSE* = 0.007, $p < .05$. Validity modulated the effects of beliefs to a greater extent in the unconstrained condition than in the constrained condition. In the former, there was a belief bias of equal size both for valid and invalid syllogisms (valid-congruent, .80; valid-incongruent, .73; invalid-congruent, .70; invalid-incongruent, .62). In the latter, belief bias was present only for the valid syllogisms (valid-congruent, .77; valid-incongruent, .65; invalid-incongruent, .53; invalid-congruent, .49). The significant three-way interaction Validity × Atmosphere × Condition shows that analytic reasoning—even though it was more prominent in the unconstrained condition—played a limited role also in the constrained condition, $F(1, 41) = 4.5$, *MSE* = 0.026, $p < .05$. In the constrained condition, valid syllogisms were less affected by atmosphere than were invalid syllogisms (valid/consistent, .87; valid/inconsistent, .54; invalid/consistent, .81; invalid/inconsistent, .21), whereas they were similarly affected in the unconstrained condition (consistent/valid, .94; consistent/invalid, .89; inconsistent/valid, .59; inconsistent/invalid, .42).

Table 6. *Mean response latencies for the eight sorts of argument in the two conditions in Experiment 2*

		Correct		*Incorrect*	
Belief		*Valid (Type 1)*	*Invalid (Type 4)*	*Valid (Type 2)*	*Invalid (Type 3)*
Congruent	Constrained task, training set	2.4 (0.5)	2.5 (0.6)	(absent)	(absent)
	Constrained task, test set	2.0 (0.6)	2.3 (0.6)	2.6 (0.7)	2.3 (0.7)
	Unconstrained task	10.9 (5.5)	13.1 (5.7)	15.9 (6.5)	15.9 (6.7)
Incongruent	Constrained task, training set	2.4 (0.6)	2.7 (0.6)	(absent)	(absent)
	Constrained task, test set	2.1 (0.6)	2.5 (0.6)	2.7 (0.7)	2.2 (0.7)
	Unconstrained task	13.9 (5.6)	16.0 (6.8)	16.0 (7.3)	16.0 (8.5)

Note: Correct/incorrect refers to the response suggested by the atmosphere. Valid/invalid denotes type of argument. Response latencies in s. Standard deviations in parentheses.

Table 6 shows the latencies of responses to the eight sorts of argument, by atmosphere, belief, and validity, in the two conditions. Latencies were analysed by a 2 × 2 × 2 × 2 mixed-design ANOVA, within-participants factors being the atmosphere consistency, validity, and belief congruence. The between-groups factor was the condition. All the main effects were significant: atmosphere, $F(1, 41) = 24.5$, $MSE = 6.23E+6$, $p < .0001$, means = 7.9 s (consistent), 9.2 s (inconsistent); validity, $F(1, 41) = 4.7$, $MSE = 4.95E+6$, $p < .05$, means = 8.3 s (valid), 8.8 s (invalid); beliefs, $F(1, 41) = 7.4$, $MSE = 7.35E+6$, $p < .01$, means = 8.1 s (congruent), 8.9 s (incongruent); condition, $F(1, 41) = 99.7$, $MSE = 1.65E+7$, $p < .00001$, means = 2.4 s (constrained condition), 14.7 s (unconstrained condition). Responses were slower in the unconstrained condition, when inconsistent with the atmosphere, for invalid syllogisms, and when incongruent with prior beliefs, fitting the corresponding effects found for the accuracy data. Atmosphere interacted reliably with belief—$F(1, 41) = 13.9$, $MSE = 3.26E+6$, $p < .001$, means = 7.1 s (atmosphere consistent, congruent beliefs), 8.6 s (atmosphere consistent, incongruent beliefs), 9.2 s (atmosphere inconsistent, congruent beliefs) 9.3 s (atmosphere inconsistent, incongruent beliefs)—which is consistent with the Atmosphere × Belief interaction in the accuracy data. In the constrained condition, responses were slower in the training set than in the test set (2.5 s vs. 2.3 s), $t(22) = 2.3$, $p < .05$, thus confirming that participants learned strategies during the training set.

Two items of the questionnaire (Items 2 and 4; Bacon et al., 2003) address the use of semantic strategies—based on the mental representation of the sets described in the premises—and four items (3, 5, 6, and 7) address the use of syntactic, linguistic strategies, allowing an estimation of how much each participant recruited different strategies (mean of the answers for each set of questions, each one on a Likert scale from 1, "not at all", to 5, "a lot").[6] The questionnaire does not distinguish between analytically correct and heuristic strategies. Participants in the time-constrained group reported using more linguistic strategies than those in the unconstrained group (3.8 vs. 3.2), $t(41) = 2.23$, $p < .05$. In contrast, participants reported more semantic strategies in the unconstrained condition than in the time-constrained condition (3.7 vs. 2.5), $t(41) = 2.93$, $p < .01$. Of the 13 participants who used paper and pencil in the unconstrained condition, 7 drew at least one diagram similar to Euler circles or connecting terms with "=", "≠", or similar symbols, suggesting that they were using some kind of semantic strategy for that syllogism, and

[6] We believe that Items 1 and 8 were not sufficiently diagnostic, and we did not use them. Instead of formulating Item 4 as a yes/no question, we required answering on a Likert scale. In our version, "5" corresponded to "a lot".

8 rewrote (at least once) the premises and/or conclusion, or used arrows to represent exchange of terms or match of quantifiers between premises and/or conclusions, suggestive of the use of mostly linguistic strategies.

Discussion

The strict time constraints and homogeneity of the syllogisms in the training set induced a prevalence of heuristic reasoning in the constrained condition, as shown by the poor performance with Type 3 syllogisms in that condition (22% correct). Performance almost doubled (43% correct) in the unconstrained condition, suggesting that participants were more likely to recruit analytical processing in that condition. The same conclusion is supported by the increased conflict between validity and use of previous beliefs in the unconstrained condition. Nonetheless, analytical processing played a minor, but significant, role in the time-constrained condition, since valid syllogisms were less affected by the atmosphere than were invalid syllogisms in that condition, and valid syllogisms were selectively affected by beliefs.

In general, the effects of atmosphere were stronger than the effects of beliefs: Consistency of atmosphere increased the proportion of correct responses by .44, whereas congruency of prior beliefs improved performance by a mere .08. The difference might have been inflated by the fact that our syllogistic conclusions were—mostly—contingently true (or false), whilst effects of beliefs are stronger where the conclusions are true (or false) by definition (Duncker, 1935; Oakhill & Johnson-Laird, 1985; Revlin et al., 1980). Nonetheless, the results show that knowledge-driven heuristics are secondary with respect to structural heuristics, at least for these sorts of problem.

Apart from atmosphere being generally stronger than beliefs, the relationships between the two heuristics are interesting. The Belief × Atmosphere interactions show that where the atmosphere suggested accepting a conclusion, it was not important whether beliefs suggested accepting or rejecting it. In contrast, where the atmosphere suggested rejecting a conclusion, true conclusions were accepted more than false conclusions. This trend can be mapped onto the previous finding that beliefs affect invalid more than valid syllogisms (Evans, Barston, & Pollard, 1983; Morley, Evans, & Handley, 2004), but in this case it applies to atmosphere instead of validity of the conclusion. On the one hand, this further corroborates the idea that atmosphere can supersede logic. On the other hand, it again suggests that belief-based heuristics are secondary with respect to other reasoning strategies. This interpretation is also suggested by the different time courses of the effects of atmosphere and beliefs. The contribution of atmosphere to responses was very strong in the short term (i.e., in the constrained condition), but decreased in the long term (i.e., in the unconstrained condition). By contrast, effects of beliefs were weak in the constrained condition—where they affected only valid syllogisms—and slightly increased in the control condition, where they affected all syllogisms. The results of the debriefing questionnaires, which showed that participants in the constrained condition used more linguistic strategies than those in the unconstrained condition, hint at the same conclusion. The conclusion that beliefs effects are secondary, both in size and in time course, is consistent with many theoretical views contending that recourse to previous knowledge does not replace other reasoning processes, but rather interferes with them—either at the response level (Oakhill & Johnson-Laird, 1985), or at the stage where premises are integrated (Cherubini et al., 1998; Klauer et al., 2000; Quayle & Ball, 2000) or represented (Revlin et al., 1980). Yet, it is inconsistent with Sloman's (1996) suggestion that beliefs must be inhibited in order to recruit other reasoning strategies, which would lead to a prediction that belief effects should be stronger in the early stages of reasoning.

GENERAL DISCUSSION

These experiments show that a variety of different styles of reasoning can be recruited by a syllogistic task and that their relative contribution to the conclusion changes with the task demands. Severe time constraints on responses induce individuals to seek and use fast heuristics. For instance, if the correct responses to the syllogisms initially presented to the participant are homogeneously consistent with the responses suggested by the atmosphere principles (Begg & Denny, 1969; Chapman & Chapman, 1959; Sells, 1936; Woodworth & Sells, 1935)—as they were in the training set of the constrained conditions of our experiments—people rapidly learn that heuristic and later generalize it even to syllogisms where the correct response is not consistent with it—as occurred in the test set of the constrained conditions. In Experiment 2, more heuristics were available: atmosphere-like heuristics, based on a logically incomplete processing of the structure of the premises, and knowledge-driven heuristics, based on beliefs. The former dominated the latter in both conditions. Furthermore, in both conditions, beliefs affected mostly syllogisms that could not be correctly solved by simple structural heuristics.

Analytic, logically consistent reasoning in strictly time-constrained tasks was present, but played a minor role. There are three pieces of evidence that indicate its contribution. First, there is the increased performance with Type 2 than with Type 3 syllogisms (Experiment 1): As explained in the Discussion of Experiment 1, it can be interpreted in multiple ways, not all of them implying correct logical processing; however, the effect can—in some interpretations—indicate a modicum of logically correct analytic reasoning affecting the time-constrained condition. Second, in the constrained condition of Experiment 2 people were less affected by atmosphere for valid than for invalid syllogisms, suggesting a residual ability to discriminate between the two. Third, beliefs affected only valid syllogisms in the constrained condition of Experiment 2 independent of atmosphere, again showing that people were at least partly able to discriminate between valid and invalid conclusions. Apart from these findings, the below-chance level performance on atmosphere-inconsistent syllogisms in the constrained conditions of both experiments and the results of the analyses of the individual patterns of responses in Experiment 1—which uncovered an almost universal bias toward the response suggested by the atmosphere—show that the contribution of analytic, logically consistent reasoning in these conditions was weak. Analytic reasoning was greater in the unconstrained conditions, where participants received an extensive preexperimental training with correctness feedback, could use paper and pencil, confronted all types of syllogism from the start, and could self-pace their responses. This is shown by the reliable increase in performance with syllogisms inconsistent with the atmosphere.

These findings help clarify dual-process views of reasoning. The style of reasoning adopted for solving a task depends upon the task demands. Analytic strategies are scarcely available where fast responses are requested. Analytic strategies are more available—even though they still are not the leading reasoning style—in less constrained contexts. Where fast responses are required, and resources are limited, people mostly recruit heuristics, preferring structural heuristics (for both abstract and thematic problems) to knowledge-driven heuristics (available for thematic problems only). Recourse to structural heuristics decreases with the relaxing of time constraints, to the benefit of analytic processing. By contrast, the contribution of beliefs does not decrease with the relaxing of the time constraints: Actually, it increases, because in the unconstrained condition beliefs affect all syllogisms, whereas in the constrained condition they affect only valid syllogisms (a possible side effect of the weakening of the atmosphere effect).

Structural heuristics depend on limited, logically incomplete processing of the structure of the premises. This limited processing can be either syntactic, considering superficial linguistic features of the premises (e.g., atmosphere or matching), or semantic, based on a representation of a logically incomplete set of models representing the set-relations between terms. This latter case is emphasized by mental models theory (MMT; Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991). The theory contends that most of the intuitive, everyday reasoning is "one-model" reasoning (Johnson-Laird, 2001), where conclusions are descriptions of a single representation of a set of possible state of affairs in which the premises are true. Typically, they are valid conclusions for easy problems, where no alternative models could be built. Yet, a one-model initial conclusion is often wrong for problems allowing for alternative models. In order to derive the correct conclusion, people must engage in effortful search for counterexamples—that is, a search for alternative models where the premises hold but the initial conclusion does not. Despite the different terminology, what in dual-process theories is known as analytic reasoning is—in mental models' parlance—a slow, logically consistent style of reasoning where people engage in the systematic search for counter examples; in contrast, heuristic reasoning is the "on the fly" reasoning where people limit themselves to considering the first model that comes to mind. Both perspectives account equally well for our data, because most of our Type 1 and Type 4 syllogisms—where the response suggested by the atmosphere was correct—require only one model, whereas most of Type 2 and Type 3—where the response was inconsistent with the atmosphere—require more than one model (Johnson-Laird & Bara, 1984). Some additional results fit nicely with the MMT view. For instance, the Beliefs × Atmosphere interaction, when analysed with respect to accuracy, showed that beliefs affected very weakly the proportion of correct responses to easy one-model problems (i.e., Types 1 and 4), whilst their contribution increased for more difficult multimodel problems (i.e., Types 2 and 3). This finding replicates previous data, and MMT offers good theoretical interpretations for it (Cherubini et al., 1998; Oakhill & Johnson-Laird, 1985; Oakhill, Johnson-Laird, & Garnham, 1989). Again, the MMT perspective can elegantly account for the unpredicted better performance observed in Experiment 1 for Type 2 versus Type 3 syllogisms: An easy, one-model heuristic working backward from the conclusion to the premises suggests the correct response for Type 2 and the wrong response for Type 3 syllogisms. On the other hand, participants' self-reports suggest that both linguistic and model-based structural heuristics were used, with linguistic strategies preferred in the constrained condition and model-based strategies preferred in the control condition. The data strongly suggest that structural heuristics were the main contributor to responding in the constrained conditions, yet they do not allow a precise estimation of how often these heuristics were linguistic (atmosphere-like), or model based.

A secondary aim of this study was to get a closer look at the cognitive processes underlying performance in the studies on the neural basis of syllogistic reasoning by Goel and colleagues (Goel et al., 2000; Goel & Dolan, 2003, 2004), which showed a partial neuroanatomical dissociation between abstract and thematic syllogisms. Committing to the dual-process view, Goel (2003) interpreted the dissociation as corresponding to a functional dissociation between analytic and heuristic processes. Our data, obtained by using the same time constraints and syllogisms as those used in those studies, suggest that an alternative is equally possible—namely, that both networks mostly refer to heuristic processes, albeit different ones: a strong activation of structural heuristics (for both abstract and thematic syllogisms), and a weak contribution of knowledge-driven heuristics (for thematic syllogisms only).

Original manuscript received 10 April 2007
Accepted revision received 17 January 2008
First published online 28 May 2008

REFERENCES

Bacon, A. M., Handley, S. J., & Newstead, S. E. (2003). Individual differences in strategies for syllogistic reasoning. *Thinking & Reasoning*, *9*, 133–161.

Begg, I., & Denny, J. P. (1969). Empirical reconciliation of atmosphere and conversion interpretations of syllogistic reasoning errors. *Journal of Experimental Psychology*, *81*, 351–354.

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

Bucciarelli, M., & Johnson-Laird, P. N. (1999). Strategies in syllogistic reasoning. *Cognitive Science*, *23*, 247–303.

Chapman, L. J., & Chapman, J. P. (1959). Atmosphere effect re-examined. *Journal of Experimental Psychology*, *58*, 220–226.

Cherubini, P., Garnham, A., Oakhill, J., & Morley, E. (1998). Can any ostrich fly?: Some new data on belief bias in syllogistic reasoning. *Cognition*, *69*, 179–218.

Cherubini, P., & Mazzocco, A. (2004). From models to rules: Mechanization of reasoning as a way to cope with cognitive overloading in combinatorial problems. *Acta Psychologica*, *116*, 223–243.

De Neys, W. (2006a). Automatic-heuristic and executive-analytic processing during reasoning: Chronometric and dual task considerations. *Quarterly Journal of Experimental Psychology*, *59*, 1070–1100.

De Neys, W. (2006b). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science*, *17*, 428–433.

Dickstein, L. S. (1978). The effect of figure on syllogistic reasoning. *Memory & Cognition*, *6*, 76–83.

Duncker, K. (1935). *Zur Psychologie des produktiven Denkens* [The psychology of productive thinking]. Berlin: Springer.

Erickson, J. R. (1974). A set analysis theory of behaviour in formal syllogistic reasoning tasks. In R. Solso (Ed.), *Loyola Symposium on Cognition* (Vol. 2). Hillsdale, NJ: Lawrence Erlbaum Associates.

Evans, J. St. B. T. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology*, *75*, 451–458.

Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, *11*, 295–306.

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning*. Hove, UK: Psychology Press.

Ford, M. (1995). Two modes of representation and problem solution in syllogistic reasoning. *Cognition*, *54*, 1–71.

Gilhooly, K. J., Logie, R. H., Wetherick, N. E., & Wynn, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. *Memory & Cognition*, *21*, 115–124.

Gilhooly, K. J., Logie, R. H., & Wynn, V. (1999). Syllogistic reasoning tasks, working memory, and skill. *European Journal of Cognitive Psychology*, *11*, 473–498.

Goel, V. (2003). Evidence for dual neural pathways for syllogistic reasoning. *Psychologia*, *32*, 301–309.

Goel, V., Buchel, C., Frith, C., & Dolan, R. J. (2000). Dissociation of mechanisms underlying syllogistic reasoning. *NeuroImage*, *12*, 504–514.

Goel, V., & Dolan, R. J. (2003). Explaining modulation of reasoning by belief. *Cognition*, *87*, B11–B22.

Goel, V., & Dolan, R. J. (2004). Differential involvement of left prefrontal cortex in inductive and deductive reasoning. *Cognition*, *93*, B109–B121.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Cambridge University Press.

Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Sciences*, *5*, 434–442.

Johnson-Laird, P. N., & Bara, B. G. (1984). Syllogistic inference. *Cognition*, *16*, 1–61.

Johnson-Laird, P. N., & Byrne, R. (1991). *Deduction*. Hove, UK: Lawrence Erlbaum Associates.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review*, *93*, 1449–1475.

Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review*, *107*, 852–884.

Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs*, *54*, 1–95.

Morgan, L. J., & Morton, J. T. (1944). The distortion of syllogistic reasoning produced by personal convictions. *Journal of Social Psychology*, *20*, 39–59.

Morley, N. J., Evans, J. St. B. T., & Handley, S. J. (2004). Belief bias and figural bias in syllogistic reasoning. *Quarterly Journal of Experimental Psychology*, *57A*, 666–692.

Newstead, S. E., Pollard, P., Evans, J. St. B. T., & Allen, J. P. (1992). The source of belief bias effects in syllogistic reasoning. *Cognition*, *45*, 257–284.

Oakhill, J., & Garnham, A. (1993). On theories of belief bias in syllogistic reasoning. *Cognition*, *46*, 87–92.

Oakhill, J., & Johnson-Laird, P. N. (1985). The effect of belief on the production of syllogistic conclusions. *Quarterly Journal of Experimental Psychology*, *37A*, 553–569.

Oakhill, J., Johnson-Laird, P. N., & Garnham, A. (1989). Believability and syllogistic reasoning. *Cognition*, *31*, 117–140.

Ormerod, T. C. (2000). Mechanisms and strategies for rephrasing. In W. Schaeken, G. De Vooght, A. Vandierendonck, & G. d'Ydewalle (Eds.), *Deductive reasoning and strategies* (pp. 131–152). Hove, UK: Lawrence Erlbaum Associates.

Ormerod, T. C., & Richardson, J. (2003). On the generation and evaluation of inferences from single premises. *Memory & Cognition*, *31*, 467–478.

Quayle, J. D., & Ball, L. J. (2000). Working memory, metacognitive uncertainty, and belief bias in syllogistic reasoning. *Quarterly Journal of Experimental Psychology*, *53A*, 1202–1223.

Revlin, R., Leirer, V. O., Yopp, H., & Yopp, R. (1980). The belief bias effect in formal reasoning: The influence of knowledge on logic. *Memory & Cognition*, *8*, 584–592.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge, MA: MIT Press.

Roberts, M. J. (2000). Individual differences in reasoning strategies: A problem to solve or an opportunity to seize?. In W. Schaeken & G. De Vooght (Eds.), *Deductive reasoning and strategies* (pp. 23–48). Mahwah, NJ: Lawrence Erlbaum Associates.

Sells, S. B. (1936). The atmosphere effect: An experimental study of reasoning. *Archives of Psychology*, *50*, 289–293.

Sloman, S. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*, 645–665.

Störring, G. (1908). Experimentelle Untersuchungen über einfache Schlussprozesse [The experimental study of syllogisms]. *Archiv für die gesamte Psychologie*, *11*, 1–27.

Wason, P. C., & Evans, J. St. B. T. (1975). Dual processes in reasoning? *Cognition*, *3*, 141–154.

Wetherick, N. E., & Gilhooly, K. J. (1995). "Atmosphere", matching, and logic in syllogistic reasoning. *Current Psychology*, *14*, 169–178.

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

Woltz, D. J., Gardner, M. K., & Bell, B. G. (2000). Negative transfer errors in sequential cognitive skills: Strong-but-wrong sequence application. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 601–605.

Woodworth, R. S., & Sells, S. B. (1935). An atmosphere effect in syllogistic reasoning. *Journal of Experimental Psychology*, *18*, 451–460.