



HAL
open science

Solving categorical syllogisms with singular premises

Guy Politzer, Hugo Mercier

► **To cite this version:**

Guy Politzer, Hugo Mercier. Solving categorical syllogisms with singular premises. *Thinking and Reasoning*, 2008, 14 (4), pp.414-453. ijn_00343123

HAL Id: ijn_00343123

https://hal.science/ijn_00343123

Submitted on 29 Nov 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

To appear in *Thinking and Reasoning*, 2008, 14(4)

Solving Categorical Syllogisms with Singular Premises

Guy Politzer

CNRS – ENS – EHESS – Institut Jean Nicod, Paris

Hugo Mercier

ENS – EHESS – CNRS – Institut Jean Nicod, Paris

Keywords: reasoning, syllogism, ecthesis

Abstract

We elaborate on the approach to syllogistic reasoning based on "case identification" (Stenning & Oberlander, 1995; Stenning & Yule, 1997). It is shown that this can be viewed as the formalisation of a method of proof that dates back to Aristotle, namely proof by exposition (*ecthesis*), and that there are traces of this method in the strategies described by a number of psychologists, from Störring (1908) to the present day. It was hypothesised that by rendering individual cases explicit in the premises the chance that reasoners engage in a proof by exposition would be enhanced, and thus performance improved. To do so, we used syllogisms with singular premises (e. g., *this X is Y*). This resulted in a uniform increase in performance as compared to performance on the associated standard syllogisms. These results cannot be explained by the main theories of syllogistic reasoning in their current state.

Solving Categorical Syllogisms with Singular Premises

The year 2008 marks the 100th anniversary of the publication by Störring of the first experimental investigation of syllogistic reasoning –possibly the first psychological investigation of reasoning *tout court*. But in spite of researchers' unabated attempts, to date there is no such a thing as a widely accepted theory that would explain lay people's performance in solving syllogisms. This paper is yet another one which concerns itself with a task that seems to be impervious to theorists' efforts and, more seriously, that might have little to teach us on real life reasoning capabilities. Let us elaborate on this latter point. We do not adopt the opinion that people never, or hardly ever, argue by use of syllogisms. On the contrary, we consider that people do exploit syllogisms, but we believe that when they do so, the arguments they actually use are less constrained than the standard textbook syllogisms, so that these should be regarded as abstractions that underlie the reasoners' arguments. There is one first -- and trivial-- sense in which this is true, namely because the classical syllogisms are formal. But there is a second and deeper reason: The classical syllogisms are idealisations that go beyond the form that is strictly observable, due to several factors such as the order of the premises and, more importantly, the absence of one of the premises in the very common enthymematic uses. What, then, are the classical syllogisms of the experimental tasks? We believe that they are exercises whose solution requires higher-level skills than do the instantiated syllogistic arguments of daily life; in fact their solution requires a metacognitive level, that is, a level that it would be necessary for individuals to reach, in case they were asked to justify their informal

argument, to formulate the associated formal argument. In brief, textbook syllogisms are abstract normative rules; the laboratory task built on these, in which a more or less explicit decision about validity is required, cannot be anything else than a metacognitive task (in the sense of testing participants' awareness of the rules that guide their own inferential production) –unless the task is meant to test the rote learning of these rules, as it used to be in former times.

The consequence of this analysis is that solving the experimental task is more difficult than understanding or producing the equivalent informal argument, and that in the absence of training in logic only skilled participants –that is, those who can spontaneously apply an effective strategy-- can cope with the syllogisms of greater logical complexity. This latter point leads us to consider the strategic aspect of reasoning, with which a number of theorists have concerned themselves in recent years (Evans, 2000; Gilhooly, 2005; Roberts, 2000; Roberts, Gilmore, & Wood, 1997; Roberts & Newton, 2005; Schaeken, De Vooght, & d'Ydewalle, 2000; Van der Henst, Yang, & Johnson-Laird, 2002). Indeed, from the beginning of the psychological investigation of syllogisms until the present time, theorists have been aware of the existence of various strategies (Bacon, Handley, & Newstead, 2003; Bucciarelli, & Johnson-Laird, 1999; Ford, 1995; Galotti, Baron, & Sabini, 1986; Marrero, & Gámez, 2004;. Matsuno, 1987; Oberlander, 1998; Stenning, & Cox, 2006; Störring, 1908), even though this fundamental fact has often been ignored. In fact, if the standard task is still worth investigating, this is not because it is directly relevant to the reasoning processes of daily life. It is

rather worth investigating because it provides information on the various strategies that people can apply at a metacognitive level.

One major question raised by Stenning, and Cox (2006) is whether participants adopt a credulous or a sceptical attitude, which leads to different interpretations of the task. Although conceptually this classification is orthogonal to the *existence* of various strategies to solve the task construed as formal (e. g., verbal or diagrammatic), we believe that participant's attitude may turn out to be correlated with (if not determining) the *use* of a specific strategy (e. g., relying more or less on heuristics).

Our conception of the syllogism task being stated, this paper will be devoted to defining with some rigour, and theoretically justify in logical terms, a strategy of major importance, of which a few experimenters have been aware to different degrees. We will show that this strategy appears in various guises but can be reduced to a single method of proof initially described by Aristotle, namely *ecthesis*. Then we will show that one of the various theoretical approaches to syllogistic reasoning turns out to be very close to the strategy in question. Finally we will show that it is possible to prime this strategy by using a non-standard task. The prediction is that whenever the strategy can be primed performance will improve, because the strategy is made available to those individuals who do not possess the skills necessary to use it spontaneously.

Aristotle's proof by ecthesis

It is well known that Aristotle used two main methods of proof. The first one is based on the identification of four "perfect" syllogisms, all in the first figure

(we keep the traditional numbering of syllogisms used by logicians). They are: *all M P, all S M / all S P*; *no M P, all S M / no S P*; *all M P, some S M / some S P*; and *no M P, some S M / some not S P*. The method consists in turning the syllogism into a perfect one by conversion or by "mutation" (changing the order of the premises). The second method is used in the two cases where the first one cannot be applied and consists in a *reductio ad impossibile*. All the cases being settled, there is no need for another method of proof. To the surprise of many commentators, Aristotle nevertheless used another type of proof, called *ecthesis* (or proof by *exposition*) . He applied it only a few times and expressed it in a rather concise manner. Here is the ecthetic proof of AAI-3 (*all M P; all M S / some S P*, usually formulated in Aristotle's writings as "*P belongs to every M; S belongs to every M / P belongs to some S*"):

"...if both P and S belong to every M, should one of the M, e.g. N, be taken, both P and S will belong to this, and thus P will belong to some S" (*Analytica Priora*, 6, 28a, Vol. 1, p. 46, Oxford translation with letters changed).

In modern times, this passage has received two different interpretations. A first construal was proposed by Lukasiewicz (1957) and Patzig (1968), who claimed that *ecthesis* amounts to assuming the existence of a category common to the subject and the predicate (or to the subject and the negated predicate) of a particular sentence. On this view the exposed entity, N, is another term (or category) of the same type as P and S. There are technical and historical difficulties with this view which nowadays has few supporters but the important

point is that a sub-category of the universal subject category is distinguished. According to the other construal supported by Lear (1980, p. 4), Mignucci (1991), Smiley (1973), Smith (1982), and Thom (1976; 1981, p. 170), ecthesis is akin to existential instantiation in natural deduction, so that the exposed entity is an individual variable, not a category. At any rate, we assume that the gist of an ecthetic proof consists in selecting a sub-category C' (that can be treated as a whole) from one of the three categories C , or an individual c from one of these in such a way that C' (or c) can be doubly predicated. This yields an entity characterized by the three terms (as C is trivially one of the predicates). The solution follows by dropping the middle term.

Here is an example of an ecthetic proof for the classical EI-3: *no M P; some M S / some S not P*. From the second premise we "expose" (i. e., extract) one or several individuals that are both M and S; because these are M, they are warranted by the first premise not to have property P; from which it follows that we are considering individuals that are S but not P, hence the solution *some S are not P*. Note that the method also applies to identify the invalid syllogisms. Consider the pair of premises close to the previous one, OI-3: *some M not P; some M S*. This time the exposed individual that is also both M and S cannot receive the predicate not-P because this predication applies particularly to some M without warrant that it applies to the one M that has been exposed: No conclusion follows.

Considered as an autonomous system, Aristotle's syllogistic has given rise to many studies by modern logicians from the viewpoint of axiomatisation. These will not concern us, with one exception, namely Thom's (1981) system which

differs from the others by two main related characteristics. First, in addition to particular and universal classical propositions, in which the subject is a category (a term variable: A = *all X are Y*; I = *some X are Y*; E = *no X are Y*; O = *some X are not Y*), it accepts singular propositions, in which the subject is an individual variable: U+ = *x is Y*, and U- = *x is not Y*. Second, it has an axiom and derived rules that specifically capture ecthesis and enable one to derive both singular and classical valid syllogisms. In brief, Thom's system includes the whole of Aristotle's syllogistic and in using it as a framework one is not studying arguments in a different domain. The set of classical propositions now enriched with singular propositions enables one to extend the classical syllogistic to incorporate singular syllogisms.

A few relevant psychological approaches

We have mentioned earlier that a few investigators have independently given their own description of a specific strategy used by some of their participants. We claim that, with hindsight, this strategy can be identified with an ecthetic proof. The cases of interest will be reviewed briefly as they are described in more detail in Politzer (2004). Consider first the Störring (1908) study, which used verbal reports. He described two main strategies, one visual (corresponding to a spontaneous use of diagrams), the other verbal, which he called the process of "insertion". In the latter case, his four participants frequently selected an end term from one premise to insert it next to the middle term in the other premise. For instance, given *some P M*; *all M S*, a participant typically said "all the M, including some P, are S" to conclude correctly *some P are S* after abstracting M.

Similarly, almost a century later, Ford (1995) observed independently that about one half of her subjects exhibited on the majority of the syllogisms a "substitution behavior" that consists in replacing one term in a premise with another. In Ford's own terms (using the same syllogism as an example) the second premise allows one to give the "value" of S to M, so that the "value" of S can be substituted for M in the first premise, leading to the conclusion.

We now consider Braine's (1998) comments on the individual differences in performance; these are attributed to differences in availability of a strategy that consists in choosing a *secondary topic*: Choose "the subset of the subject of which the middle term can or cannot be predicated ("the S that are, or are not, M," as determined by the premise relating S and M)" (p. 321). Once the secondary topic is chosen, it is transferred into the other premise. Here is how Braine described the solution of *no M P; all M S / some S not P*: The second premise provides the secondary topic, *the S that are M*; from the first premise, the S that are M are not P by modus ponens, hence *some S are not P* by existential generalisation. In brief, in Braine's analysis while the proof is *primed by an instantiation*, the completion of the proof requires an existential generalisation when the conclusion is particular, and a universal generalisation when it is universal. In summary, it is clear that the three descriptions are variants of ecthetic proofs, even though the investigators do not explicitly underscore the crucial role of the selection of a subcategory or an individual. By contrast, this is at the heart of the last theoretical approach that we are going to consider, namely Stenning's notion of *individual cases*.

Stenning and Yule (1997) have proposed that syllogisms exist and are soluble owing to one of their structural properties which they call "identification of individual cases". An individual can be characterized by the fact that it possesses, or does not possess, the properties that define the three categories S, M, and P, so that there are eight types of individual: S+M+P+, S+M+P-, S+M-P+, S+M-P-, S-M+P+, S-M+P-, S-M-P+, and S-M-P-. For each syllogism, the joint premises warrant or do not warrant the existence of such individuals and in the affirmative the syllogism has a conclusion. The authors describe two algorithms for identifying the individual cases that constitute the conclusion, one analogical (graphical) and the other sentential (by rules), and they claim that these constitute two variants of a common underlying abstract individual identification algorithm. The first one (Stenning & Oberlander, 1995) uses a variant of Euler circles to represent each premise mood (A, E, I, and O). Among the regions defined by the two circles, there is a "critical region" for each mood, that is, a region that must exist if the premise is true (as opposed to regions that are only consistent with its truth). For example, for *all X are Y* the critical region is the part common to X and Y while the region made of the Y that are not X is only optional. When the premises are combined, a region critical for one premise need not be so for the other premise, so that there are two cases: If there remains no common critical region, the syllogism has no valid conclusion; if critical regions survive, the syllogism has a conclusion defined by these regions applied to the end terms. Interestingly, the authors distinguish a critical region by using a mark (a cross), which illustrates that the existence of a conclusion depends on the permanency of

the existence of a part of a category (or the individual that instantiates it, in actual fact, the "exposed" entity) through the combination of the premises.

The other algorithm has three parts. The first part consists in the choice of a "source premise", that is, a premise that provides its categories as the first two terms of the tentative individual description (one of these being necessarily the middle term). The second part is an attempt to complete the description with the second end term and here the quality of the middle term in both premises is considered. If the qualities match and M is the subject of the other premise, a modus ponens is applied (using the conditional sentence associated to the premise) whose conclusion (which is the predicate of the second premise) provides the third term of the individual description; if the qualities do not match, and M is the predicate of the other premise, this means that there is an M+ and an M-, which allows a modus tollens whose conclusion (which is the subject of the other premise) again provides the third term of the description; if none of the two previous cases occurs, there is no conclusion. In the third part, the M term is deleted and a quantity is attributed. From the present viewpoint, the first part is crucial in that it starts the exposition. Take for example AO-2, *all P M, some S not-M*: The source is the second premise, which contains a negated M (M-); this provides S+M- for the first two terms of the individual description whose survival in relation with P will be decided on the next two steps of the algorithm. Since in the first premise M is predicate and affirmative, modus tollens applies to *if P then M, not-M*, yielding *not-P*, that is, P-, which completes the description into S+M-P-, hence the conclusion *some S not-P*. It is also apparent that the sentential

algorithm details Störring's insertion and abstraction process, that the first part specifies Braine's choice of a secondary topic, and that the second and third steps coincide with Ford's substitution process (and application of rules). In sum, from the inception of the psychological investigation of syllogisms different experimenters have made observations that are consistent with the exposition strategy, but these observations were rather vague and incidental. Stenning and colleagues have rediscovered this strategy and turned it into a full-fledged theory.

The two algorithms presented by Stenning and Oberlander (1995) are descriptions which reveal that the essence of the syllogisms and of their solution at the computational level consists in the inquiry about the possibility that an individual initially characterised as a function of S and M, or P and M, be eventually characterised as a function of S and P. Stenning and Yule (1997) presented an experiment in which they used a novel task, the "individual task". Participants were invited to try to determine, for all pairs of premises, whether there was any type of individual that must exist and, in the affirmative, to describe it (as an S that is an M and a P, or an S that is an M and not a P, etc.) This offers the possibility to make fine-grained predictions, in particular with regard to figural effects, which were well supported by the data. However, this task differs from the standard syllogism task by the absence of a conclusion formulated as a quantified sentence. In the present study, we will stay closer to the standard task. Any procedure or modification of syllogisms that suggests to the reasoner the consideration of one or several individuals that have both properties defined by the subject and predicate of a premise (S and M, or P and M), that is, the

exposition method, will act as a trigger to the use of the Störing-Ford "insertion/substitution" strategy, which should lead to the solution of the valid syllogisms. One way of realising this is using the singular syllogisms of Thom's system. In effect, our basic hypothesis is that singular sentences naturally orientate the reasoner towards individuals, which in turn should facilitate the solution.

Experiment

Method

Materials

Special care was taken with respect to the grammatical and logical form of the premises.

The grammatical form of the premises. Our aim was to study the effect on syllogism solving that may result from the replacement of a classical particular (plural) premise (I or O) by a singular premise that invites the reasoner to focus on an individual. To do so, the demonstrative "this" suggests itself. However, this choice introduces a possible confounding factor because "this" is not only singular but also definite, whereas "some" is plural and indefinite. (It could be objected that "some" can be singular; however, this is only the limiting case of a vague quantifier as any positive number from 1 to the cardinal of the set of reference satisfies "some". In addition, in French, the language of the experiment, "some" is rendered by "certains", which is plural). Consequently, in addition to the standard condition using the classical premises I and O, two control conditions were introduced, one using the definite plural demonstrative "these" (French "ces"), and

the other using the indefinite singular determiner "there is a" (French "il y a un"). This had the other advantage that it provided a second singular expression to test our hypothesis. "This" and "there is a" differ in that the demonstrative adds reference to an individual designated by ostension or by previous mention (Corblin, 1987). In an artificial setting (a test or an experiment) there is an element of pretense by which "this" is understood *as if* the designation had actually occurred. (They further differ pragmatically: Whereas "there is a" is paraphrasable by "at least one", and therefore synonymous with "some", it seems to readily trigger the conversational implicature "no more than one", which "this" does not do). Table 1 displays the four conditions together with the type of premise which defines them.

Table 1. The four conditions as defined by the grammatical form of the premises.

	Plural (control conditions)	Singular (experimental conditions)
Indefinite	Condition "some" (classical premise) I: some X are Y O: some X are not Y	Condition "there is a" there is an x that is a Y there is an x that is not a Y
Definite	Condition "these" these X are Y these X are not Y	Condition "this" this x is a Y this x is not a Y

The logical form of the premises. The valid syllogisms were defined following Thom's (1982) axiomatisation. There are 10 pairs of premises with a valid conclusion. Eight of them have one universal premise (A or E) and one non

classical premise (singular in Thom's system, singular or demonstrative in the present case, e. g. , *all MP; this SM*) and the last two pairs were such that both premises are non-classical. To have a greater number of valid syllogisms, the "dual" forms of these pairs were also used (that is, the pairs in the reverse order which yield a conclusion in the opposite direction and which, with one exception, Thom did not consider). This yielded a total of 16 "valid" pairs of premises. With respect to the traditional syllogistic, it should be kept in mind that only 48 pairs of premises are relevant to the present investigation, namely those that have at least one particular premise; these can be partitioned into two subsets: One subset consists of the 16 pairs that are the counterpart of those just defined (e. g. , *all MP; some SM* is the classical counterpart of *all MP; this SM*); the other subset consists of the remaining 32 "invalid" pairs. To have an equal number of valid and invalid syllogisms, only sixteen pairs were selected out of these and were given one non-classical premise (half with a universal premise and the other half with a particular premise). In brief, there were 32 pairs of premises, 16 with a valid conclusion and 16 without a valid conclusion. They appear in Table 2 (in which U designates the non-classical moods) together with the conclusion that follows validly when applicable. In the classical ("some") condition, the conclusions are the ones that are classically admitted when both orders are allowed for the end terms. Regarding the non-classical conditions, two points deserve attention.

The first point concerns the mood in which the conclusion is expected to be given, classical (I or O) or non-classical. We consider both conclusions as valid for the following reasons. One, a demonstrative conclusion ("this", "these")

entails a "some" conclusion (by existential generalisation). Two, we assume a "some" conclusion to be consistent with a demonstrative conclusion by existential instantiation, even though the demonstrative noun phrase is not a variable but is rather akin to a constant. This is assumed to obtain, in the same way that participants who accept the element of pretense built in the task actually accept to treat "this m" as unspecified in a sentence such as "this m is a P", that is, as a variable. In brief, the difference between answers in I vs. U+, (and O vs. U-) will not be taken into account to determine the correctness of the answers but only as a qualitative indicator of participants' inferential process.

The second point regards the syllogisms numbered 15 and 16 in Table 1. Although they are invalid in the two conditions that use indefinites ("some" and "there is a"), it is acceptable to consider them as valid in the two conditions that use definites ("this" and "these"). This stems from the fact that in the classic interpretation to which we stick, in the case of indefinites the subject refers to two different groups of individuals (in the case of "some") or to two different individuals (in the case of "there is a"), whereas in the case of definites, the subjects of the two premises can be understood as referring to the same individual(s). This can be illustrated with syllogism 15. In the "there is a" condition, the premises are: "there is an m that is a P ; there is an m that is an S". The assumption here is that there are two "m", one of which "P" is predicated, and the other of which "S" is predicated. In contrast, in the "this" condition, the premises are: "this m is a P / this m is an S". It seems natural to assume conversationally that it is the same "m" that "P" and "S" are predicated of

(although this is by no means logically obliged). These two assumptions were checked during a preliminary inquiry and were further corroborated by the answers given by the participants.

The content of the premises. Sixteen brief scenarios depicting plausible situations were introduced by a short sentence. They were varied so that there were four scenarios referring to each of the following types: Animals, vegetables, artifacts, and professions. The scenarios were made specific enough to eliminate potential belief biases linked to the associated general scripts. Here is an example followed by the two premises (translated from French): *At a conference, all the organisers are Asians; this organiser is a jurist.*

The problems were presented in booklets containing eight pairs of premises. Classically (e.g. Dickstein, 1978; Johnson-Laird & Bara, 1984; Johnson-Laird & Steedman, 1978), the participants are given a whole set of 64 pairs of premises, and it is possible that tiredness and the use of simple heuristics can explain some of the answers that are produced. The chances of such processes occurring here are reduced by the small amount of problems to be solved. The format offered to the participants for their answers is novel. The participants had to choose between “there is a necessarily true conclusion” and “there is no necessarily true conclusion”. If they chose the first option, they had to fill up one of two lines that differed by the order of the end terms by choosing among six sentence types (the four classical moods and the two non-classical moods). Here is an example (which completes the previous example):

there is a necessarily true conclusion:

.....Asian(s).....jurist(s)

.....jurist(s).....Asian(s)

there is no necessarily true conclusion

The conclusion had to be chosen among the following sentence types:

All the.....are.....; No.....is.....; Some.....are.....; Some.....are not.....;

This.....is.....; This.....is not..... For the conditions with "these" and "there is a",

the last two sentence types were: *These.....are.....; These.....are not.....;* and

There is a.....that is...; There is a.....that is not..., respectively.

For each condition, four booklets were created, each containing eight problems determined as follows. The set of 32 problems was partitioned into four subsets of eight problems in a pseudo-random manner, that is, with the constraint that the number of valid and invalid problems was approximately balanced (either 4 of each or 3 valid and 5 invalid). This determined four types of booklet. Within each of these four types, four versions were created that differed by the position of the problems, providing a control for a possible order effect. . Then, each pair of premises was given a content, so that no booklet had the same content twice.

The booklets that contained the problems were preceded with a booklet that explained the task using two examples (the premises leading to AA-4 and to AA-2, for a valid and an invalid syllogism, respectively) and contained the six response types participants were allowed to use.

Participants and procedure

The participants were three hundred and thirteen undergraduate students. Most of them were majoring in psychology and all were native speakers of

French. They were recruited in a faculty library and were working at their own pace.

Design

The design was between-participants. Each participant was randomly allocated to one of the four conditions.

Predictions

The predictions are straightforward. The main prediction is that on both singular forms (*this* and *there is a*) performance will be higher than on the classical condition (I and O sentences) because the singular forms suggest the exposition strategy (applied to an individual). There are two other predictions. For the definite (demonstrative) sentences, performance should be higher in the singular than in the plural form because the singular is again more likely to suggest the exposition strategy. A subsidiary prediction is that performance may be improved (as compared to the classical condition) in the plural demonstrative condition (*these*) because there is a possibility that the designation communicated by the demonstrative primes the exposition process (applied to the part of a category).

Results

Comparison between conditions

Table 2 presents the mean percentages of correct answers for each syllogism in each condition. Collapsed across syllogisms these percentages were equal to 45.9 % for the classical ("some") condition and to 55.6%, 59.8%, and 63.9% for the "these", "there is a", and "this" conditions, respectively, showing a

clear difference in performance between the classical and the non classical conditions.

A statistical analysis that enables one to consider the various syllogisms individually was preferred to averaging over problems. Indeed, it is important to check whether there are specific syllogisms at the source of any difference. Pairwise comparisons between conditions were made by calculating for each syllogism the direction of the difference in the percentages of correct answers. A sign test was then performed on the sign of the difference.

The three non classical conditions resulted in significant improvements with respect to the classical "some" condition: There was an improvement on 28 of the 32 differences ($p < .0001$) for "there is a", on 26 of the 30 differences ($p < .0001$) for "this", and on 21 of the 30 differences ($p < .05$) for "these" (Recall that there are only 30 relevant comparisons for the demonstratives). The effectiveness of the use of a singular or a demonstrative determiner (as opposed to the particular) is perhaps best highlighted by the following observation: Out of 32 problems, in 25 cases performance with "some" was the lowest of the determiners, whereas there are only two cases where it was the highest. In brief, the effect of the manipulation was very systematic and it was not due to some special subset of syllogisms.

Table 2. Percentage of correct answers for each condition and each syllogism.

(The figures of the syllogisms are numbered according to the logicians' traditional numbering).

U = the non classical moods, A, I, E, O the classical moods (all, some, no, some...not)

+ = an affirmative sentence, - = a negative sentence.

* nvc = no valid conclusion ° nvc when U is indefinite

Syllogism No, premises mood, conclusion and figure	Condition			
	some	these	there is a	this
1: A U+ / U+ 1	70.6	55.6	100.0	90.0
2: U+ A / U+ 4	76.5	62.5	89.5	75.0
3: E U+ / U- 1	44.4	44.4	60.0	57.9
4: U+ E / U- 4	25.0	40.0	45.0	56.3
5: E U+ / U- 2	29.4	33.3	36.8	50.0
6: U+ E / U- 2	11.1	33.3	30.0	44.4
7: A U- / U- 2	22.2	37.5	35.0	45.0
8: U- A / U- 2	11.8	40.0	36.8	56.3
9: A U+ / U+ 3	42.9	63.2	83.3	66.7
10: U+ A / U+ 3	58.8	56.3	83.3	64.7
11: E U+ / U- 3	33.3	25.0	40.0	52.6
12: U+ E / U- 3	11.8	45.0	35.0	56.3
13: U- A / U- 3	13.3	21.1	44.4	28.6
14: A U- / U- 3	27.8	44.4	45.0	44.4
15: U+U+ / U+° 3	66.7	75.0	65.0	69.2
16: U- U+ / U-° 3	62.5	73.7	65.0	68.8

(Table continues)

Syllogism No, premises mood, conclusion and figure	some	these	there is a	this
17: A U- / nvc* 1	26.7	70.0	55.0	73.2
18: U- A / nvc 4	40.0	64.7	40.0	65.0
19: I U+ / nvc 1	47.1	47.1	78.9	70.6
20: U+ I / nvc 4	93.8	50.0	88.9	71.4
21: O U- / nvc 4	68.8	90.0	80.0	82.4
22: U- O / nvc 4	57.9	81.3	65.0	80.0
23: O U+ / nvc 1	43.8	50.0	70.0	77.8
24: U+ O / nvc 4	57.1	52.6	78.9	40.0
25: I U- / nvc 1	52.9	65.0	60.0	70.6
26: U- I / nvc 4	52.9	68.8	55.0	70.0
27: E U- / nvc 1	58.8	76.5	60.0	72.2
28: U- E / nvc 4	62.5	80.0	73.7	85.7
29: A U+ / nvc 2	13.3	44.4	36.8	55.6
30: U+ A / nvc 2	38.9	47.1	45.0	61.1
31: E U- / nvc 2	61.1	66.7	75.0	66.7
32: U- E / nvc 2	82.4	78.9	72.2	80.0
mean	45.9	55.7	60.3	63.9

The only other significant difference concerns the singular vs plural demonstratives: Performance was higher for "this" than for "these" in 24 out of 32 cases ($p < .01$). The difference between the two singular determiners showed a trend for "this" to produce higher performance than "there is a" close to significance ($p = .06$). In brief, the singular vs plural manipulation was clearly effective, whether applied to the definite ("this" vs "these") or the indefinite ("there is a" vs "some") wording; and the definite vs indefinite manipulation did affect the plural forms ("these" vs "some") as expected (with a trend for the singular forms).

Discussion

To prime the exposition strategy, the selection of a specific individual from a premise was made more available by using (i) a singular expression ("there is a") that focuses on a single individual; (ii) a demonstrative ("these") that identifies individuals; or even better, (iii) an expression that is both singular and demonstrative ("this"). (As recalled earlier, classical logicians used either a subset or an individual in theorising ecthesis). It was predicted that an increase in performance would follow, which was strongly and unambiguously confirmed. The fact that the use of the two singular determiners was effective, and the use of the plural demonstrative was effective too (although to a lesser extent) indicates that what was crucial in the manipulation is the identification of one or even several *individuals*. We will now add some more support to our approach by having a closer look at the figural effect, then we will consider the generality of

our approach with regard to other quantifiers before turning to alternative explanations of the results.

On the figural effect

One of the most robust effects in the experimental literature on syllogistic reasoning concerns the order of the terms in the conclusion: It has been dubbed the "figural effect". In its more standard formulation (Johnson-Laird & Bara, 1984; Johnson-Laird & Steedman, 1978), it states that participants tend to favour an SP conclusion for figure 1 (MP; SM) and a PS conclusion for figure 4 (PM; MS). To test for this effect in our data, we need to identify the relevant responses: They are the correct conclusions that can be converted and the erroneous categorical conclusions because in both cases it is not logic that constrains the reasoner's choice between the SP and PS orders. This leaves 14 syllogisms (1, 2, and 17 to 28) in whose conclusions we will examine the order of the terms, SP vs PS. All four conditions collapsed, we obtain 64 comparisons between the two possible orders; 45 go in the direction predicted by the figural effect, 12 in the opposite direction (there are seven ties). This is statistically highly significant ($z = 4.28, p < 3.10^{-5}$) and we conclude that the figural effect was replicated in the present experiment.

The theory of case identification suggests an explanation for the figural effect, namely the process of exposition which tends to maintain the exposed entity as the final subject (this is at work in the "source founding" of the verbal algorithm of Stenning and Yule). An initial attempt to expose a term has to start with the selection of a subject end-term. Consequently, if an attempt at applying

the exposition strategy fails at a later stage, it is likely that a trace of its precursor will subsist, namely the upholding of the subject of one premise as the conclusion's subject. Therefore, it can be expected that the figural effect should be at least as strong in the experimental conditions as it was in the standard condition. Indeed, in each of the three non-classical conditions the effect was statistically significant but fell short of significance in the "some" condition: Comparing the SP versus PS answers, the data are the following (sign tests were used for comparisons). For "this" ten vs three, and three ties, $p < .05$; for "there is a", thirteen vs one, and two ties, $p < .001$; for "these" twelve vs three, and one tie, $p < .02$. In the "some" condition, the figures are ten vs five, and one tie ($p > .15$) showing that the well-established trend was present but with numerical values that suggest a smaller size of the figural effect for "some" than for the non-classical determiners.

In fact, we are in a position to refine the analysis by focusing on the second figure (PM, SM) and make a novel prediction: The end term that appears in the conclusion as a subject will tend to coincide with the one that has a demonstrative or a singular determiner (as opposed to universal) because this is the one that is identified as the term (category) to be exposed. Compare, for instance, the following two pairs in the second figure: *no P M; this S M* and *this P M; no S M*. We expect a tendency to choose S in the former case and P in the latter as the conclusion's subject because these coincide with the terms initially exposed, should the reasoner engage in the exposition process. Furthermore, we expect that this process does not carry over to "some" because this determiner has

little or no potential to prime the exposition; and this explains why no figural effect has been reported in the literature for the second figure. The prediction was tested by considering the eight problems in the second figure and counting the conclusions (with the same criteria as used earlier) of the type SP and PS. Collapsing over the three non-classical determiners, the 24 comparisons yielded fifteen differences in the predicted direction, six in the opposite direction and six ties; this differs significantly from chance (sign test, $p < .05$). As expected, the same comparisons applied in the "some" case failed to show any effect (there were three differences in the predicted direction, two in the opposite direction and three ties).

Syllogistic reasoning with other quantifiers

Although the logical tradition since Aristotle has limited the study of categorical syllogisms to four quantified sentences, other quantifiers are used in natural language, and generalised syllogistic systems have been proposed (e. g., Peterson, 2000). Oaksford and Chater (1999) rightly point out the advantage, in terms of generality, for a model as theirs to include nonclassical quantifiers in addition to the classical ones. This nicely applies to *ecthesis* too: The exposition strategy functions also with syllogisms whose premises use quantifiers such as *few* or *most*. For instance, consider *all M P; most S M*. Pick up one of these S that are M; from the major premise, this SM must be a P, hence an S that is a P, and because this can be reiterated to most S, we get *most S are P*. One can contrast this with *most M P; all S M*. From the minor premise, picking up any single S we get an SM; now because the major premise is not universal, it does not attribute P

to just any kind of M and it is unknown whether our individual SM qualifies to be a P or not, so that nothing follows. In sum, syllogisms with generalised quantifiers fall under the exposition strategy, giving this approach a wide scope. Finally, it should be noted that historically Aristotle used the proof by exposition mainly as a method of proof for modal syllogisms, which indicates that it can be applied beyond the domain of categorical syllogisms.

Alternative explanations

What remains to be discussed is the possibility of alternative explanations for our results, in the framework of the theories of syllogistic reasoning.

One well-known approach can be ruled out right away: The "atmosphere effect" applied to the present materials would imply that participants had a tendency to repeat the determiner in the conclusion. If this were the case, it should apply identically to all determiners, including "some" and no difference between conditions should exist. This does not mean that no participant ever used such a heuristic, but this strategy cannot explain the results.

Next, the approach to syllogistic reasoning based on monotonicity (entailingness) properties of quantifiers (Geurts, 2003; for an introduction, see Westerståhl, 2001) can accommodate singular and demonstrative sentences: These have monotonicity properties, the same ones in fact as their particular counterpart, that is, "this", "these", and "there is a" are left and right monotone increasing with an affirmative sentence and they are left increasing and right decreasing with a negative sentence, just like "some". But this identity is at the same time very embarrassing for this approach because identical monotonicity

properties implies identical treatment and therefore identical performance, contrary to the present results.

Turning now to the probabilistic model of Chater and Oaksford (1999), we face the problem of how to integrate the three determiners in the theory, that is, to define the semantics of the three sentences *this X is Y*, *these X are Y*, and *there is an X that is Y* within the probabilistic framework (this concerns the computational level of the theory). Recall that the probability of "no", "all" and "some" is defined by considering a reference set X, and the sentence *det X are Y* is attributed a probability value by considering $p(Y/X)$, so that "no" has probability 0, "all" has probability 1 and "some" has a positive probability. This semantics entails for "this" and "there is a" the same probability as "some": In effect, knowing that *this X is Y* (or *there is an X that is Y*), the probability that any given X is Y is comprised between 0 (excluded) and 1 as at least one X is Y (namely the X being designated) and there is no information whether other X are Y. *These X are Y* has a slightly higher probability $p(Y/X)$ than *this X is Y* because the plural indicates at least two. We will not further discuss this semantic interpretation; we just intend to show that it is in principle possible to accommodate the three determiners in the theory. Turning now to the algorithmic level, it is crucial to determine the relative informativeness of the determiners as the working of the heuristics depends on this. Clearly, because "all" entails "this/these/there is a", (e. g., *all the jurists are Asians* entails *these jurists are Asians*) which in turn entail "some" (and similarly for the negative scale, "no"

entails "this/these/there is a...not", which entail "some...not"), the informativeness scales are the following:

$$I(\text{all}) > I(\text{this/these/there is a}) > I(\text{some}), \text{ and}$$

$$I(\text{no}) > I(\text{this/these/there is a}) > I(\text{some...not})$$

It can be seen that the three determiners are in the same relation of informativeness with respect to the universal quantifier as "some" is.

Consequently, the operation of the heuristics which is based on the identification of the least informative premise (the "min-premise") must yield the same result:

Whenever "some" is the min-premise, "this"/"these"/"there is a" will be too and all the conclusions are identical, *mutatis mutandis*, that is, changing "some" for "this/these/there is a". This means that no difference in performance is to be expected between the classical syllogisms and the non-classical ones. One could invoke that the O-heuristic (avoid O conclusions) does not carry over to "this"/"these"/"there is a"; but even if this could be justified, this applies to only 5 out of 32 problems (those with valid O conclusions). In sum, the probabilistic model seems unable to account for the present results. Note that the ecthetic approach, which supports the notion that what is at the basis of syllogism solving (at least in the laboratory task), namely the determinate characterisation of individuals in an all-or-none manner, is wholly incompatible with the probabilistic view.

Finally, we consider the mental model theory (Bucciarelli & Johnson-Laird, 1999). What could be the models for the three determiners? It seems that

there are two solutions. The first one is to add a "tag" to one token (for "this" and "there is a") or to two tokens (for "these") as follows:

this X is Y	$\rightarrow x$	y	these X are Y	$\rightarrow x$	y
		x		$\rightarrow x$	y
					x
					y
this X is not Y	$\rightarrow x$	-y	these X are not Y	$\rightarrow x$	-y
		x		$\rightarrow x$	-y
					y
					y

The other premise has no tag (except in two problems), so that after combination of the two premises the models have not changed, except for the tag(s). If anything, this makes the problems harder because of the tag(s), so there is no explanation for the results. But there is another solution which, in taking into account the pragmatic nuances described earlier, may be considered more faithful to the interpretation of the determiners, even though in mental model theory there is no special device in the models to take into consideration the implicatures that affect the particular quantifier. Another reason to be cautious is that the implicatures do not seem to be triggered in the syllogism task (Roberts, Newstead & Griggs, 2001). With these reservations in mind, the modification of the models for "some" and "some...not" to represent the three determiners is straightforward: One needs to consider only one token (one line) for the singular determiners (and two tokens for the plural) so that the models become:

this X is Y	x	y	these X are Y	x	y
				x	y
this X is not Y	x	-y	these X are not Y	x	-y
	x	-y		x	-y

For the valid syllogisms the correct conclusion is obtained straightforwardly. Take for instance EU+3: *no M is P; this M is S / this S is not P*. The model for *no M is P* is the standard one:

[m] -p
[m] -p
[p]
[p]

to which we integrate the model for *this M is S*, which yields:

[m] -p s
[m] -p
[p]
[p]

from the first line of which one can read off *this S is not P*. Notice that because there are no other tokens for (s, m) to integrate, there is only one model. One can verify that the modified models provide the conclusion of all the valid syllogisms which we have considered. Remarkably, the solution, which is always particular or singular (or duplicated in case of "these") appears on one line of the model in the form of an "individual type".

For the invalid syllogisms, an inspection case by case indicates that either no full description of an individual exists in at least one model, or a full description exists in two models but they are different: In brief, no common conclusion (full description of an individual type) can be read off the models, which means that the syllogism is invalid. The improvement in performance cannot be attributed to the smaller number of models, but to the smaller number of lines necessary to build the models. However, even if this account seems reasonable, the match between the erroneous conclusions licensed by the models and the conclusions observed is very poor. We conclude that a modified version of mental model theory could account for the main trend in the results (although not for the various moods of the answers). This is in agreement with Stenning and Oberlander (1995) who showed that the mental model theory can be reinterpreted within their registration diagram framework. In sum, the modified models can be viewed as a method that makes appear the type of individual on a single line when it exists and shows its non-existence in the invalid case. Note that the difficulty to reconcile the detailed machinery of the standard mental model theory with the exposition strategy was pointed out by Stenning and Yule (1997). The predictions for the "individuals task" based on mental models were not supported by the data. Clearly people who apply exposition keep the mid-term in its initial position, contrary to the mental model hypothesis of various reordering operations that involve the mid-term.

Considering the failure of the various theories to explain our data, it is clear that any claim made by any one of these to apply to all syllogisms and to all

the individuals is untenable. It is to that extent that they are seriously challenged. One may argue that none of them makes such a sweeping claim. In that case, the challenge still exists, but it is limited to the demonstration that the reasoning of a sizeable proportion of individuals escapes them, so limiting their scope and explanatory power.

To conclude, logical solving of categorical syllogisms in the laboratory by non technical means (i. e., excluding methods such as premise conversion, the use of predicate logic or diagrams) requires the application of an adequate strategy, namely the exposition strategy that aims to the identification of the individual type that characterises the conclusion. We have hypothesised and experimentally confirmed that if this strategy is primed by use of an appropriate determiner, performance improves significantly: Less skilled reasoners who otherwise would have failed for lack of a spontaneous use of the exposition strategy now have greater chance to apply it. The reason why even with this help performance remains far from perfect is that the solution is cognitively demanding as it requires several steps such as, in some cases, the execution of a modus tollens (see the individual identification algorithm of Stenning & Oberlander, 1995): Once on the right way, there are still obstacles to get over.

In this experiment we have not carried out an individual analysis of participants' responses. The present study constitutes but a first step showing that on the average performance can be enhanced. As suggested by the literature reviewed earlier, an unknown proportion of participants can engage in the exposition strategy on an unknown proportion of syllogisms. What we did is

increase this proportion by prompting the use of this strategy. The next step will be to study the possibility to help reasoners initially identified as users of inappropriate strategies (mainly heuristics that may or may not be successful) to shift to the use of ecthesis and get better chance of success. A related direction of research, if we are right in assuming that our manipulation cued to the exposition strategy those individuals who would not have used it spontaneously, is to look for an interaction between the effectiveness of the manipulation and a measure of cognitive capacity -- individuals with lower capacity would benefit relatively more from the manipulation.

Finally, a related question is how the exposition strategy fits in with the dual-process accounts of reasoning (Evans, 2007; Sloman, 1996; Stanovich, 2004). Here it is useful to consider the means by which ecthesis is activated. The spontaneous use of the strategy in the standard task clearly seems to belong to the explicit system 2. However, when the use of the strategy is triggered by the mere presence of a singular determiner, the exposition of an individual is offered automatically and free of effort by the meaning of the determiner and this concerns system 1. Of course, one cannot exclude that after solving a few problems some individuals learn the exposition method and then become able to apply it more consciously, that is, the method would come under the control of system 2. It would be interesting to see whether there is a transfer effect from singular to standard syllogisms.

References

- Bacon, A. M., Handley, S. J., & Newstead, S. E. (2003). Individual differences in strategies for syllogistic reasoning. *Thinking and Reasoning, 9*, 133-168.
- Barnes, J. (Ed.) (1984). *The complete works of Aristotle. The revised Oxford translation*. Princeton: Princeton University Press.
- Braine, M. D. S. (1998). Steps toward a mental-predicate logic. In M. D. S. Braine & D. P. O'Brien (Eds.), *Mental logic* (pp. 273-331). Mahwah, NJ: Erlbaum.
- Bucciarelli, M., & Johnson-Laird, P. N. (1999). Strategies in syllogistic reasoning. *Cognitive Science, 23*, 247-303.
- Chater, N., & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology, 38*, 191-258.
- Corblin, F. (1987). *Indéfini, défini et démonstratif*. Geneva: Droz. [Indefinite, definite and demonstrative].
- Dickstein, L. S. (1978). The effect of figure on syllogistic reasoning. *Memory and Cognition, 6*, 76-83.
- Evans, J. St. B. T. (2000). What could and could not be a strategy in reasoning. In W. Schaeken, G. De Vooght, G. d'Ydewalle (Eds.), *Deductive reasoning and strategies* (pp. 1-22). Mahwah, NJ: Erlbaum.
- Evans, J. St. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgment*. Hove: Psychology Press.
- Ford, M. (1995). Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition, 54*, 1-71.

- Galotti, K. M., Baron, J., & Sabini, J. P. (1986). Individual differences in syllogistic reasoning: Deduction rules or mental models? *Journal of Experimental Psychology: General*, *115*, 16-25.
- Geurts, B. (2003). Reasoning with quantifiers. *Cognition*, *86*, 223-251.
- Gilhooly, K. J. (2005). Working memory and strategies in reasoning. In M. J. Roberts & E. J. Newton (Eds.), *Methods of thought. Individual differences in reasoning strategies* (pp. 57-80). Hove: Psychology Press.
- Johnson-Laird, P. N., & Bara, B. G. (1984). Syllogistic inference. *Cognition*, *16*, 1-61.
- Johnson-Laird, P. N., & Steedman, M. (1978). The psychology of syllogisms. *Cognitive Psychology*, *10*, 64-99.
- Lear, J. (1980) *Aristotle and logical theory*. Cambridge: Cambridge University Press.
- Lukasiewicz, J. (1957). *Aristotle's syllogistic from the standpoint of modern formal logic*. Oxford: Clarendon Press.
- Marrero, H., & Gámez, E. (2004). Content and strategy in syllogistic reasoning. *Canadian Journal of Experimental Psychology*, *58*, 168-180.
- Matsuno, T. (1987). Cognitive style and representational strategies in categorical syllogistic reasoning. *Tohoku Psychologica Folia*, *46*, 97-102.
- Mignucci, M. (1991). Expository proofs in Aristotle's syllogistic. In H. Blumenthal & H. Robinson (Eds.), *Aristotle and the later tradition* (pp. 9-28). Oxford Studies in Ancient philosophy. Oxford: Clarendon Press.

- Oberlander, J. (1998). Abstraction and individual style in syllogistic and heterogeneous reasoning. *Kognitionswissenschaft*, 7, 33-38.
- Patzig, G. (1968) *Aristotle's theory of the syllogism*, Dordrecht: Reidel. [Original German edn, 1958].
- Peterson, P. L. (2000). *Intermediate quantifiers*. Aldershot, U.K.: Ashgate.
- Politzer, G. (2004). Some precursors of current theories of syllogistic reasoning. In K. Manktelow & M.-C. Chung (Eds.), *Psychology of reasoning. Theoretical and historical perspectives* (pp. 214-240). Hove: Psychology 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, G. d'Ydewalle (Eds.), *Deductive reasoning and strategies* (pp. 23-48). Mahwah, NJ: Erlbaum.
- Roberts, M. J., Gilmore, D. J., & Wood, D. J. (1997). Individual differences and strategy selection in reasoning. *British Journal of Psychology*, 88, 473-492.
- Roberts, M. J., Newstead, S. E., & Griggs, R. A. (2001). Quantifier interpretation and syllogistic reasoning. *Thinking and Reasoning*, 7, 173-204.
- Roberts, M., & Newton, M. J. (2005). *Methods of thought. Individual differences in reasoning strategies*. Hove: Psychology Press.
- Schaeken, W., De Vooght, G., & d'Ydewalle, G. (Eds.). (2000). *Deductive reasoning and strategies*. Mahwah, NJ: Erlbaum.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119(1), 3-22.

- Smiley, T. J. (1973). What is a syllogism? *Journal of Philosophical Logic*, 2, 136-154.
- Smith, R. (1982). What is Aristotelian ecthesis? *History and Philosophy of Logic*, 3, 113-127.
- Stanovich, K. E. (2004). *The Robot's Rebellion*. Chicago: Chicago University Press.
- Stenning, K., & Cox, R. (2006). Reconnecting interpretation to reasoning through individual differences. *Quarterly Journal of Experimental Psychology*, 59, 1454-1483.
- Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science*, 19, 97-140.
- Stenning, K., & Yule, P. (1997). Image and language in human reasoning: A syllogistic illustration. *Cognitive Psychology*, 34, 109-159.
- Störring, G. (1908). Experimentelle untersuchungen über einfache Schlussprozesse. *Archiv für die Gesante Psychologie*, 11, 1-127.
[Experimental research on simple inferential processes].
- Thom, P. (1976). Ecthesis. *Logique et Analyse*, 74-75-76, 299-310.
- Thom, P. (1981) *The syllogism*. München: Philosophia Verlag.
- Van der Henst, J.-B., Yang, Y. , & Johnson-Laird, P. N. (2002). Strategies in sentential reasoning. *Cognitive Science*, 26, 425-468.
- Westerståhl, D. (2001). Quantifiers. In L. Goble (Ed.) *The Blackwell guide to Philosophical logic* (pp. 437-460). Oxford: Blackwell.