THINKING AND REASONING, 2002, *8* (1), 1–20

When is a conclusion worth deriving? A relevance-based analysis of indeterminate relational problems

Jean-Baptiste Van der Henst
Institut Jean Nicod, Paris, France and University of Leuven, Belgium

Dan Sperber
Institut Jean Nicod, Paris, France

Guy Politzer
CNRS, Saint-Denis, France

When is a conclusion worth deriving? We claim that a conclusion is worth deriving to the extent that it is relevant in the sense of relevance theory (Sperber & Wilson, 1995). To support this hypothesis, we experiment with "indeterminate relational problems" where we ask participants what, if anything, follows from premises such as *A is taller than B*, *A is taller than C*. With such problems, the indeterminate response that nothing follows is common, and we explain why. We distinguish several types of determinate conclusions and show that their rate is a function of their relevance. We argue that by appropriately changing the formulation of the premises, the relevance of determinate conclusions can be increased, and the rate of indeterminate responses thereby reduced. We contrast these relevance-based predictions with predictions based on linguistic congruence.

From any set of premises, an infinity of deductively valid conclusions can be derived. For instance, from the single premise *The cat is on the mat*, it follows, by the rules of propositional calculus, that *It is not the case that the cat is not on the mat* or that *The cat is on the mat or the dog is in the kitchen*. However, in the

Correspondence should be addressed to Jean-Baptiste Van der Henst, K.U Leuven, Laboratory of Experimental Psychology, 102 Tiensestraat, 3000 Leuven, Belgium. Email: jvanderhenst@hotmail.com

We thank Richard Griggs, Maxwell Roberts, Carlos Santamaría, Walter Schaeken, Jean Baratgin, Ira Noveck, and Deirdre Wilson for their helpful comments on earlier versions of this article. We also thank Jean-Marc Bernard for his help in statistical treatment. Jean-Baptiste Van der Henst is supported by a Marie Curie fellowship.

© 2002 Psychology Press Ltd
http://www.tandf.co.uk/journals/pp/13546783.html DOI.10.1080/1354678014300007 1

psychology of reasoning, there are many problems where participants, when asked what, if anything, follows from some set of premises, are judged to have given the right answer when they say that nothing follows. Of course, neither the psychologist nor the participants are demonstrating illogicality in such cases. Instead, it is tacitly understood that the participants are being asked for a valid conclusion *worth deriving* (rather than for just any logically valid conclusion). But what exactly is the content of this tacit understanding? What makes a conclusion worth deriving? Here we propose an answer to this question: a conclusion worth deriving is one that is *relevant* in the sense of relevance theory (Sperber & Wilson, 1995). We test this hypothesis experimentally by analysing participants' performance in indeterminate relational problems.

DETERMINATE AND INDETERMINATE RELATIONAL PROBLEMS

Some of the inferences most investigated by researchers on reasoning are *relational* inferences (Burt, 1919; Byrne & Johnson-Laird, 1989; Carreiras & Santamaría, 1997; Clark, 1969a, b; DeSoto, London, & Handel, 1965; Hunter, 1957; Huttenlocher; 1968; Newstead, Manktelow, & Evans, 1982; Piaget, 1921; Potts, 1972; Quinton & Fellows, 1975; Roberts, 2000); Trabasso, Riley, & Wilson, 1975. So-called *three-term series problems* well illustrate this type of inference, for instance:

Bill is better than Pete,
Pete is better than John.

Typically, participants have to infer a new relation not explicitly mentioned in the premises, which enables them to answer a question asked by the experimenter (What is the relation between Bill and John? Who is the best?). Three-term series problems have been investigated in many studies (for a review, see Evans, Newstead, & Byrne, 1993, Ch. 6) and the central issue has been how the premises are mentally represented. This has given rise to a debate between advocates of the *analogical* and *linguistic* approaches. According to the analogical approach (DeSoto et al., 1965; Huttenlocher, 1968), information conveyed by the premises is integrated in the form of a "spatial array" or "mental model" (Byrne & Johnson-Laird, 1989; Johnson-Laird, 1983). Supporters of the linguistic approach, whose main advocate has been H.H. Clark (1969a, b), claim instead that the representation of a premise is linguistic and basically corresponds to the abstract proposition describing a relation between a predicate and one or more entities (e.g. BETTER [Bill, Pete]).

Psychologists have also investigated problems in which one or more relations were left indeterminate. For example the premises,

Bill is better than Pete
Bill is better than John,

are traditionally said to be *indeterminate* because the relation between Pete and John is left unspecified. The question typically asked by experimenters may require knowledge of the relation between the two unrepeated terms, as in "Who is the worst?" (this question does not have a determinate answer) or not, as in "Who is the best? (this question has a determinate answer).

Broadly, experimental results show that it is more difficult to deal with indeterminate problems than with determinate ones (Byrne & Johnson-Laird, 1989; Hayes-Roth & Hayes-Roth, 1975; Mani & Johnson-Laird, 1982; Moeser & Tarrant, 1977; Warner & Griggs, 1980). According to supporters of the analogical approach, this is because it is impossible to build a unique integrated representation from an indeterminate set of premises (Byrne & Johnson-Laird, 1989).

Indeterminate and determinate problems have essentially been approached in the same way. The "representation question" has been the crucial one, and evidence from the study of the two types of problem has been taken to bear on exactly the same issues. The main difference between them has been seen as one of relative difficulty. On the other hand, the willingness of participants to spontaneously draw a conclusion from such problems has never been investigated. For this, the question should not be, for instance, "Is Pete better than John?" or "Who is the best?", but simply "What, if anything, follows from the premises?".

The question "What if anything follows?" has indefinitely many logically valid answers, but none of them stands out as particularly compelling. Although no previous tests had been conducted, it was predictable that many participants would say that nothing follows. On the other hand, when participants do give positive responses, these provide clear evidence of how they interpret the task and what they see as a conclusion worth deriving.

The importance of participants' interpretations of a psychological task, and the influence of pragmatic factors on this interpretation, are increasingly recognised in the psychology of reasoning (e.g., Evans, 1995; Hilton, 1995; Noveck, 2001; Politzer & Macchi, 2000; Politzer & Noveck, 1991; Sperber, Cara & Girotto, 1995; Thompson, 2000; Van der Henst, 1999, 2000). Here, we want to show how, at least with indeterminate relational problems, a pragmatic approach helps to predict and explain participants' responses. In particular, we attempt to alter the rate and specific contents of positive responses by manipulating the formulation of the problem. We are not directly addressing the more standard issues of how participants' mentally represent the premises of relational problems, or what procedure they follow in drawing a conclusion. The issue we are addressing is that of participants' expectations and goals.

COGNITIVE EXPECTATIONS, RELEVANCE, AND THE (IN)DETERMINACY OF THE PREMISES

The difference between determinate and indeterminate three-term problems is that, in indeterminate cases, it is impossible to infer a relation between two of the terms that was not already described in the premises. This may violate participants' expectations about the kind of conclusion they should be able to draw in the context of a reasoning task. But what exactly are these expectations? What, in general, makes a conclusion worth inferring? What kind of conclusion do participants expect the experimenter to expect them to draw? We claim that conclusions worth deriving are conclusions that are *relevant.*

What is relevance?

In relevance theory, relevance is seen as a property of inputs to cognitive processes (e.g., stimuli, utterances, mental representations). An input is relevant to an individual at a certain time if processing this input yields *cognitive effects*. Examples of cognitive effects are the revision of previous beliefs, or the derivation of *contextual conclusions*, that is, conclusions that follow from the input taken together with previously available information. Everything else being equal, the greater the cognitive effects achieved by processing an input, the greater its relevance. On the other hand, the greater the effort involved in processing an input, the lower its relevance. Everything else being equal, it is clearly conducive to greater cognitive efficiency to aim at greater relevance in the inputs one processes. (The assumption that human cognition is geared to maximising relevance is called the "cognitive principle of relevance".)

Let us illustrate. Suppose it is already known that only the youngest of Barbara, Pamela, and Jane has to stay at home. In this context, the set of premises {*Barbara is older than Pamela*, *Pamela is older than Jane*} is relevant because it yields the three contextual conclusions *Jane has to stay at home*, *Barbara does not have to stay at home*, *Pamela does not have to stay at home.* This first set of premises is *more* relevant in this context than another set {*Barbara is older than Pamela*, *Barbara is older than Jane*}, which yields only one of the three previous contextual conclusions: *Barbara does not have to stay at home*. On the other hand, the first set of premises is *less* relevant in the context than the single premise *Jane is the youngest*, which yields the same three contextual conclusions with less processing effort.

The relevance of a piece of information is relevance to its user. In trying to produce relevant information for the use of an audience, or for oneself in the future, the effort to be minimised is that of the user, and this may typically involve some extra effort on the part of the producer of information. For instance, a speaker who knows that both Barbara and Pamela are older than Jane might make the effort of reformulating this knowledge, and say "Jane is the youngest", so as to produce an utterance that is optimally relevant to the audience in the context.

Relevance theory does not invoke an absolute measure of mental effort or of cognitive effect, and it does not assume that such a measure is available to the spontaneous workings of the mind. What is assumed is that, not always but quite often, the actual or expected relevance of two inputs can be compared. Such comparisons help individuals allocate their cognitive resources. They also make it possible to manipulate the relevance factor in experimental research.

How is relevance involved in reasoning experiments?

Relevance is involved in reasoning experiments in three ways:

1. As in any act of communication, communicated information is automatically presented as relevant to the addressee (this is known, in relevance theory, as the "communicative principle of relevance"). In the experimental case, this means that participants are expected to treat the set of premises communicated to them by the experimenter as if they were *relevant to them* (generally with an element of pretence, as when reading fiction or engaging in pretend play, as participants have no real use for the pseudo-information that, say, Bill is better than Pete). The information may be immediately relevant to them if they have background knowledge in the context of which it carries cognitive effects. It may also be potentially relevant, in that it may provide premises that would be useful in future contingencies.
2. Just as in any ongoing discourse, where each new utterance is expected to be relevant in a context partly determined by the interpretation of the previous utterances, each new premise in a problem is expected to be *relevant in the context of the other premises*.
3. Participants expect their response to be *relevant to the experimenter*. These responses could not achieve relevance to the experimenter by informing her of a solution that she already knows, but they can by informing her about the participants' abilities, in particular their ability to derive conclusions relevant to themselves. Relevance to the experimenter is achieved in the same way as in answering an exam question (the examiner knows the answer; what he does not know, and what is relevant to him, is whether the examinee knows or is able to work out the answer).

How does extracting conclusions from a set of premises contribute to its relevance?

When participants are presented with a specific question (e.g., "Which of Pete, Bill and John is the best?"), the premises are typically relevant to them by allowing them to deduce the requested answer. However, when participants are asked "What, if anything, follows?", it is less clear why, and in what way, they should go beyond restating the premises. Their answer might demonstrate that they have understood the potential relevance of the premises for further reasoning, but then a puzzle arises. How can deductively deriving a conclusion

and adding it to, or substituting it for, an initial set of premises yield a more relevant point of departure for further reasoning, given that nothing can be derived from this conclusion that wasn't already derivable from the initial premises?

Here is the answer. A set of premises with some deductively derived conclusion added cannot be more relevant than the initial set *on the effect side*, but it can be more relevant *on the effort side*. If the initial set of premises is expected to be relevant in some context where it carries certain cognitive effects, then the augmented set (or, in some cases, just the conclusion) might carry the same effects in that context, but at a lower cost in terms of effort. In fact, the prior deduction of some specific conclusion may be a preliminary and effort-costly necessary step towards deriving these cognitive effects. In that case, adding the deductive conclusion to (or, in some cases, substituting it for) the initial set of premises increases its expected relevance. It can be seen as preparatory work that improves the individual's readiness to make use of the premises in the context of further information.

We frequently encounter information that we think is likely to prove useful in the future. We then retain this information, and often process it in such a way as to optimise its potential usefulness. Suppose, for instance, that you arrive in a holiday resort where you plan to spend a month with your family. You learn that there are three doctors in the resort, Smith, Jones, and Williams. You also learn the following two pieces of information: {*Smith is a better doctor than Jones*, *Jones is a better doctor than Williams*}. At the time, you don't need a doctor, but you might in the future, and would then want to visit the best doctor in town. So the information is potentially relevant to you. You might just store the two pieces of information, but from a cognitive point of view it would be more efficient to draw the conclusion: *Smith is the best doctor* straight away. By drawing this conclusion now, you prepare for future circumstances in which you would need a doctor. By adding this conclusion to the two initial premises, you are left with a set of premises for future inference with a greater expected relevance, as its exploitation will require fewer inferential steps. Moreover, if you expect not to need information about the other two doctors, it may be sufficient to remember just the conclusion *Smith is the best doctor*, replacing the initial two-premise set with the single derived conclusion, thus reducing the memory load.

What relevant conclusion should participants expect to be able to infer?

Expecting each premise to be relevant in the context of the others means expecting some conclusion to be derivable from the premises taken together that was not derivable from the premises taken separately. Combining expectations of relevance to oneself (point 1 described earlier) with the expectation that the premises will be relevant to one another (point 2), participants should expect to be able to derive from the premises taken together some conclusion that

increases the relevance of the information contained in these premises by reducing the effort needed to achieve future cognitive effects. Moreover, combining this with the expectation of producing a response that will be relevant to the experimenter (point 3), participants should expect to be able to produce a response that demonstrates to the experimenter their ability to derive from the premises taken together a conclusion potentially relevant to themselves in the way just explained.

EXPERIMENT

To investigate participants' conception of a conclusion worth deriving, we asked what, if anything, followed from simple three-term relational problems. We compared their responses on the one hand to determinate vs indeterminate problems, and on the other hand, and in much greater detail, to two versions of indeterminate problems.

Types of possible conclusions and their relevance

Whether or not participants' produce a conclusion depends, we have claimed, on the expected relevance of the conclusions they can draw. In the case of determinate problems, such as {*A is taller than B*, *B is taller than C*}, the conclusion that can be drawn about the relationship between the two unrepeated terms, A and C in our example, stands out as clear and potentially relevant. We therefore expect most participants to produce such conclusions.

In the case of indeterminate problems, no obviously relevant conclusion stands out, and so we must look at the different types of possible conclusions. We take as an example the set of premises {*A is taller than B*, *A is taller than C*}. We can distinguish three types, with the third dividing into two sub-types:

(1) Conclusions based on a single premise, for instance:
A is taller than some other item
A is taller than B or 2 + 2 = 4
If B is taller than A, then A is shorter than B
(2) Conclusions connecting the two premises, for instance:
A is taller than B and A is taller than C
A is taller than B or A is taller than C
(3) Conclusions integrating the two premises:
(3a) Single-subject conclusions (i.e., with the repeated term A as subject), for instance:
A is taller than B and C
A is the tallest
(3b) Double-subject conclusions (i.e., with the conjunction of the unrepeated terms B and C as subject), for instance:
B and C are shorter than A
B and C are the shortest

Everyone would presumably predict that participants are unlikely to produce conclusions of type (1), but it may still be worth spelling out why. It cannot be just that these conclusions contain no new information, as this is true of all deductive conclusions. It cannot be that these conclusions are immediately obvious. Many (for instance: *If B is taller than A*, *then A is shorter than B*) are not. We justify this prediction by noting that conclusions of type (1), which are based on only one premise, violate the expectation that both premises are relevant, and that each is relevant in the context of the other.

Everybody would presumably predict that participants are unlikely to produce conclusions of type (2), but again, why? To begin with, we would argue that processing effort spent in conjoining (and even more, disjoining) two given premises in a conclusion is mis-directed. Typically, conjoined propositions have to be separated into two atomic propositions which can then serve, jointly or separately, as premises for further inferences. By contrast, there are few occasions on which it is useful to conjoin or disjoin premises in order to perform some real-life inference (for an argument that this is never necessary, and that inferences based on introduction rules are trivial and irrelevant, see Sperber & Wilson 1995, pp. 95–103). In any case, even if some relevance could be found for conclusions of type (2), those of type (3) are definitely more relevant, and we have argued that participants try to produce the most relevant conclusion possible.

Single-subject conclusions are more relevant than double-subject conclusions

With indeterminate problems, conclusions of type (3) are potentially relevant. Given a context of independently obtained information, they might serve as useful inputs for further inferences. There are significant differences between conclusions of type (3a) (single-subject) and (3b) (double-subject) which affect their expected relevance. Both types come in either comparative form (*A is taller than B and C*, *B and C are shorter than A*) or superlative form (*A is the tallest*, *B and C are the shortest*). In fact, as the superlative form must be interpreted with respect to the set of three items A, B, and C—that is, *the tallest* or *the shortest* means *the tallest of A*, *B*, *and C* or *the shortest of A*, *B*, *and C*—these conclusions, whether of type (3a) or (3b), and whether of comparative or superlative form, have the same informational content. However, this content is not presented in the same way in conclusions of types (3a) and (3b). In particular, in one case the grammatical subject and topic consists of a single item (A), and in the other, it consists of the conjunction of two items (B and C). In other terms, the information is "about" A in one case, and "about" B and C in the other. It is a commonplace that the linguistic structure of a statement indicates what the statement is about and affects the way it is attended, processed, and remembered.

Everything else being equal (and, in particular, the informational content being the same), single-topic conclusions have a greater expected relevance than multiple-topic conclusions. This has to do with the way new information is likely to be obtained and processed in human cognition generally. Our identification of individual items and categories is not random. It tends to pick out autonomous things or sets of things that have relatively stable interlocking properties and coherent and predictable behaviour. This tends to maximise the inferable consequences of recognising something as a given item or as belonging to a given category.

Human categorisation is geared to representing information in a way that maximises the cognitive effects that can be derived from it, and minimises the effort needed to derive these effects; in other words, it is geared to the maximisation of relevance (this, by the way, is an application of the cognitive principle of relevance to the issue of categorisation). For instance, it is quite generally more useful from a cognitive point of view to individuate as entities spatio-temporally continuous and autonomous objects (e.g., the dog Fido) rather than discontinuous spatial or temporal parts of objects (e.g., the entity made up of the dog Fido and the cat Julius). Similarly, it is more cognitively useful to group in the same category entities with similar and interrelated properties rather than less coherent ensembles (e.g., a category of cats is more useful than a category of animals or vehicles whose name in English begins with a "d").

Thus, more often than not, we pick out, we store, we use as premises, and we communicate information about single items or categories. Of course, almost any piece of information relates several items or categories, but in verbally or mentally representing the information, we typically treat it as information about a specific item or category, for instance by making this the subject or topic of an utterance. When we are presented with information about two or more items, we are likely to break it down for storage and further processing. For instance, if you learn that *John and Billy are friends of Martha*, you are likely to remember and use in the future as distinct pieces of information that *John is a friend of Martha*, that *Billy is a friend of Martha*, and/or that *Martha has two friends*, *John and Billy*. Information is stored under the mental "entries", "concepts", or "files" that we have for recognised entities or categories. You are much more likely to have one mental entry for John and one for Billy, than one for both John and Billy (unless they together make up a category such as the Blues Brothers). In ordinary circumstances, this is an optimal way of storing information, as you are more likely, in the future, to obtain new information about any given individual than about any given pair of individuals.

Given these considerations, conclusions of type (3a), with a single subject and topic, are potentially more relevant than those of type (3b), with a double subject and topic. Again, there is no difference in actual informational content. The difference is only in the form, which affects the effort needed to store, retrieve, and use the conclusion in a context of further information. The

difference in relevance between the two types of conclusions is strictly on the effort side.

Conclusions of types (1) and (2) are not relevant at all. Conclusions of type (3a) are more relevant than those of type (3b). We therefore predict, first, that participants who do give a determinate answer will produce only conclusions of type (3), and, second, that they will produce more single-subject conclusions (type 3a) than double-subject conclusions (type 3b). This second prediction is far from obvious, especially if it is extended from the specific type of indeterminate problem we have used as an example to another, symmetric, kind of indeterminate problem. As we will show, a possible and, at first blush, reasonable approach based on Clark's (1969a,b) *principle of congruence* would make the same predictions as ours for one type of problem and exactly opposite predictions for the other.

Let us contrast, then, two types of indeterminate problems. In one type, the two premises of the indeterminate problem have the same subject, as in the example we have already used {*A is taller than B*, *A is taller than C*}. In a second type, the two premises have different subjects, as in {*B is taller than A*, *C is taller than A*}.[1]

With both same-subject premises and different-subject premises, it is possible to derive single-subject or double-subject conclusions. Let us represent the four possibilities in table form (see Table 1). Our prediction applies not just to same-subject premises, but also, for the same reasons, and in the same way, to different-subject premises. We predict that, faced with different-subject premises, participants will produce single-subject conclusions, such as *A is shorter than B and C* or *A is the shortest* more frequently than double-subject conclusions such as *B and C are taller than A*, or *B and C are the tallest*.

Clark (1969a,b) studied tasks where participants had to answer a question (e.g., *Who is the shortest*?) on the basis of the premises of a relational problem (e.g., {*A is shorter than B*, *A is shorter than C*}). He argued that participants' performance was best predicted by invoking a *principle of congruence*, which claims that in answering, participants' are influenced by the linguistic form of the premises and the question. If the adjective used in the question (e.g., *Who is the* ***tallest****?*) is the same as that used in the premises (e.g., {*A is* ***taller*** *than B*, *A is* ***taller*** *than C*} then participants answer the question faster. When the adjective in the question (*Who is the* ***tallest****?*) is different from that used in the premises ({*B is* ***shorter*** *than A*, *C is* ***shorter*** *than A*}), then participants answer more slowly. This congruence principle might be justified in terms of processing effort: when the adjectives are not congruent, an extra step of conversion has to take place in order to verify the putative conclusion.

---

[1]These two types of problems, with the same adjective in the two premises, are called "homogeneous" indeterminate relational problems, to be contrasted with "heterogeneous" problems such as {*A is taller than B, C is shorter than A*}, with different adjectives in the two premises. We are not considering heterogeneous problems in this article.

TABLE 1
Two types of determinate conclusions with same-subject and different-subject premises

	Same-subject premises *A is taller than B* *A is taller than C*	*Different-subject premises* *B is taller than A* *C is taller than A*
Single-subject conclusions	*A is taller than B and C* *A is the tallest* (Predicted to be more frequent than double-subject conclusions by both the congruence and the relevance approaches)	*A is shorter than B and C* *A is the shortest* (Predicted to be more frequent than double-subject conclusion conclusions by the relevance approach)
Double-subject conclusions	*B and C are shorter than A* *B and C are the shortest* (Predicted to be less frequent than single-subject conclusions by both the congruence and the relevance approaches)	*B and C are taller than A* *B and C are the tallest* (Predicted to be more frequent than single-subject conclusions by the congruence approach)

Extending this principle of congruence from the question-answering tasks studied by Clark to the production task we are considering here, one might predict that participants are more likely to produce a conclusion with an adjective "congruent" with the adjectives used in the premises (we are not, of course, attributing such an extension of his principle to Clark himself). Again, this hypothesis could be defended in terms of processing effort. It is less costly to use an adjective already used in the premises and highly accessible in working memory than to replace it by its opposite. Therefore, from same-subject premises {*A is taller than B*, *A is taller than C*}, participants should derive a conclusion such as *A is taller than B and C*, or *A is the tallest*. From different-subject premises {*B is taller than A*, *C is taller than A*} they should derive a conclusion such as *B and C are taller than A*, or *B and C are the tallest*.

Another linguistic "congruence" argument for the same prediction could be made on syntactic rather than lexical grounds. To avoid unnecessary processing effort, the syntactic form of the conclusion should be congruent with that of the premises. Thus, with same-subject premises {*A is taller than B*, *A is taller than C*}, A is the subject and participants should be inclined to produce conclusions with A as subject, i.e., *A is taller than B and C*, or *A is the tallest*. On the other hand, with different-subject premises {*B is taller than A*, *C is taller than A*}, B and C are in subject position and participants should be inclined to produce conclusions with B and C as subjects, i.e., *B and C are taller than A*, or *B and C are the tallest*.

So the linguistic approach we are considering, based on considerations of lexical and/or syntactic congruence between the premises and the conclusion, makes a prediction identical to ours for problems with same-subject premises, but exactly the opposite prediction for problems with different-subject premises.

Single-subject conclusions are more relevant when derived from different-subject premises than from same-subject premises

Relevance theory suggests yet another prediction. With different-subject premises, single-subject conclusions involve not only an integration of the premises, but also a linguistic reformulation, both syntactic and lexical. From a syntactic point of view, term A, which is a constituent of the predicate in the premises, becomes the subject in the conclusion, while B and C, which are subjects of the two separate premises, become constituents of the predicate (at least when the conclusion is in the comparative form). From a lexical point of view, the adjective *taller* in the premises is replaced by its opposite *shorter* in the conclusion (or *shortest* in the superlative form). With same-subject premises, on the other hand, single-subject conclusions keep the syntax and lexicon of the premises intact. Thus, more effort is required to formulate single-subject conclusions with different-subject premises than with same-subject premises.

We have claimed that a conclusion deduced from a set of premises achieves relevance when it enables contextual implications to be derived, in some present or future context, with less effort than would be needed to derive these implications directly from the premises and the context. What makes a conclusion relevant is not the amount of effort involved in producing it, but the saving on future effort needed to use it in further inference. Often, the producer of a conclusion has to make some extra effort in order to reduce the user's effort (whether the user is another person or the producer him/herself at a later time). With same-subject premises, a single-subject conclusion is almost effortlessly accessible; so deducing it on the spot leads to little reduction in the effort needed to derive contextual implications in a future context. By contrast, with different-subject premises, deriving a single-subject conclusion involves some genuine effort, and thus reduces the effort needed to derive future contextual conclusions. Hence, a single-subject conclusion is more relevant when derived from different-subject premises than from same-subject premises. Moreover, participants may feel that conclusions produced by such an appropriate increase of effort are relevant to the experimenter, in that they give evidence of participants' reasoning ability.

We therefore predict that single-subject conclusions will be produced more often with different-subject premises than with same-subject premises. If we are right to claim that participants are more likely to give determinate answers to

indeterminate problems when they see these answers as more relevant, we should observe a higher rate of determinate answers with different-subject premises than with same-subject premises.

Predictions

To recapitulate our predictions:

1. There will be fewer *nothing follows* answers with determinate relational problems than with indeterminate ones.
2. All or nearly all determinate responses will be conclusions integrating the two premises.
3. With both same-subject premises and different-subject premises, there will be more single-subject conclusions than double-subject conclusions.
4. There will be relatively more single-subject conclusions with different-subject premises than with same-subject premises. As a consequence, there will be fewer indeterminate answers.

Method

Participants. A total of 139 French-speaking students from the University of Paris 13, ranging in age from 17 to 26, took part in this experiment.

Materials. In order to avoid training effects (see Quinton & Fellows, 1975), participants received only one determinate problem and one indeterminate problem. Four types of determinate problems were used: {*A is taller than B*, *B is taller than C*}, {*B is taller than C*, *A is taller than B*}, { *A is taller than B*, *C is shorter than B*}, {*C is shorter than B*, *A is taller than B*}. Two types of indeterminate problems were used:

Same-subject premises:	*A is taller than B*
	A is taller than C
Different-subject premises:	*B is taller than A*
	C is taller than A

In all problems the letters were replaced by French one-syllable first names such as Luc, Paul, Frank. Participants were given a sheet of paper that included the instructions and the two problems, each followed by the word "response". Half the participants received the indeterminate problem first, and the other half received the determinate problem first. All the material was in French.

Design. All participants received one determinate problem and one indeterminate problem so that they acted as their own controls for this comparison. Each participant received an indeterminate problem that had either

same-subject premises or different-subject premises, so that there were two experimental groups for this comparison.

Procedure. Participants were tested in two groups in their classroom. Their written instructions were as follows:

> In this task, you will be presented with two short reasoning problems. Each of these problems consists of two propositions. The two problems are completely independent. You have to solve them one after the other and not at the same time. Your task consists first in reading the two propositions. After that, if you think that you can *deduce* a conclusion from these two propositions, then write down the conclusion you have deduced after the word *response.* On the other hand, if you think that you cannot deduce *anything* from these two propositions, then write down "nothing can be deduced" after the word *response.* When you have finished with one problem, you can move to the next one.

Results and discussion

Classification of responses. The responses were classified into several categories and sub-categories. Responses such as "nothing can be deduced" were labelled *indeterminate* answers.

The *determinate* correct answers were divided into the following sub-categories:

- Responses with "A" as grammatical subject (single-subject responses). They can be comparative (A is taller/shorter than B and C) or superlative (A is the tallest/shortest, A is the tallest/shortest of the three).
- Responses with "B and C" as grammatical subject (double-subject responses). They could be comparative (B and C are taller/shorter than A) or superlative (B and C are the tallest/shortest), but in fact they were all comparative.
- Responses expressing a conjunction of responses of the former two types of response (A is taller than B and C; B and C are shorter than A).
- Responses based on a single premise.
- Finally, there were a few erratic responses.

Test of the predictions. For the first prediction, we consider the rate of indeterminate answers. Overall, it was smaller (8.2%) with the determinate problems than with the indeterminate ones (44.8%), as predicted; this obtained with a between-subjects comparison for problems presented in the first position (8.3% vs 37.9%: $\chi^2 = 15.1$, $p < .001$) as well as for problems presented in the second position (8.0% vs 51.5%: $\chi^2 = 28.7$, $p < .001$). A within-subject comparison confirmed this result: there were 55 participants who gave an indeterminate response to an indeterminate problem and a determinate response

to a determinate problem, whereas only 6 participants gave an indeterminate response to a determinate problem and a determinate response to an indeterminate problem, a very highly significant difference (McNemar test, $\chi^2 = 37.7, p < 10^{-7}$). In brief, in accord with our first prediction, participants were much less inclined to express indeterminate "nothing follows" conclusions with determinate problems than with indeterminate ones. The great majority of the determinate answers to the determinate problems consisted in stating the relationship between the two unrepeated terms (74.1% of the participants expressed a correct relation between A and C).

Our other three predictions concern only indeterminate problems, on which we now focus. The percentages of answers to the two indeterminate problems are presented in Table 2. All but one of the correct determinate answers to the indeterminate problems integrated the two premises. There was no answer that merely connected the two premises and only one participant produced a response based on a single premise. This confirms the second prediction.

To test the third and fourth predictions, we need to extract, respectively, two columns and two rows from the contingency table. The analysis can be appropriately performed using Bayesian statistics. To test the third prediction, we first consider both types of indeterminate problems (same-subject premises and different-subject premises) together. More determinate answers were expressed with a single grammatical subject "A" (34.5% of the total number of responses) than with a double grammatical subject "B and C" (14.4% of the total number of responses). This result was submitted to a Bayesian analysis for frequency comparisons (Bernard, 1998).[2] It indicated that the probability of the single-subject conclusion being more frequent than the double-subject conclusion was equal to .999, a credibility level that strongly supports the third

TABLE 2
Distributions (%) of the types of conclusions drawn from indeterminate relational problems with same-subject premises and different-subject premises

	Same-subject premises: *A is taller than B* *A is taller than C*	*Different-subject premises:* *B is taller than A* *C is taller than A*	*All problems*
"Nothing follows"	54.0	30.8	43.2
Single-subject conclusions	25.7	44.6	34.5
Double-subject conclusions	13.5	15.4	14.4
Other determinate conclusions*	6.8	9.2	7.9
	100	100	100

*Includes conjunctions of single-subject and double-subject conclusions and errors (fewer than 3%).

prediction. As can be seen in Table 2, this result is obtained separately for same-subject premises and different-subject premises. In particular, for different-subject premises ({B is taller than A, C is taller than A}), there is an even greater difference in percentages between single-subject conclusions (44.6%) and double-subject conclusions (15.4%). This difference confirms our prediction and disconfirms the competing prediction based on the linguistic approach, which went in the opposite direction. Although this result might be seen as posing a problem for the linguistic approach and not for the analogical approach, it does not by itself imply that the premises are represented analogically. It merely implies that the linguistic representation of the premises, if there is one, does not determine the conclusion that participants are willing to derive.

We now turn to the fourth prediction, about the rate of single-subject conclusions derived from same-subject or from different-subject premises. We argued that the greater the relevance of a conclusion, the greater the likelihood that participants will produce it. In particular, as single-subject conclusions are more relevant in the context of different-subject premises than in the context of same-subject premises, they should be more frequently produced in the former context than in the latter. This is what was observed (44.6% vs 25.7%). A Bayesian analysis based on these percentages indicated a level of credibility of .985 for different-subject premises being more frequent than same-subject premises, which strongly supports the fourth prediction. As a consequence of this increase as single-subject responses to different-subject premises, the rate of indeterminate conclusions was predicted to decrease (as the choice of a double-subject conclusion should not be affected by the premise type). This is indeed what happened (30.8% vs 54.0%).[3] Contrary to the prediction of the congruence approach, and as predicted by the relevance approach, there were even more (absolutely as well as relatively) single-subject than double-subject conclusions derived from different-subject premises.

GENERAL DISCUSSION

This experiment illustrates one of the roles of pragmatic factors in experimental reasoning tasks, and shows that these factors can themselves be studied experimentally. Although the existence, and indeed pervasiveness of pragmatic

[2] Computation was based on BAYCAT, version 1.01 (Bernard, 1998).

[3]There is a possible objection to the result supporting the fourth prediction. It might be objected that the increase in single-subject responses to different-subject premises could be due to a bias induced by the materials. Recall that half the participants received a determinate problem before the indeterminate one. In the first (determinate problem) some premises were of the type *A is taller than B; C is shorter than B*. Participants who subsequently received an indeterminate problem with different-subject premises, such as *B is taller than A*; *C is taller than A*, could have been prompted to use *shorter* (mentioned in the first problem), thus artefactually producing a single-subject response. This possibility was examined. There were 12 participants who gave a single-subject response and had received the indeterminate problem in second position. Only three of the twelve had a first problem capable of prompting a *shorter* response, and one did so.

factors is generally acknowledged thanks, in particular, to the work of Paul Grice (1989), the nature and effect of these factors is rarely considered in any detail. In fact, these neglected factors affect not only participants' performance, but also the thinking of experimenters.

Consider, for instance, our first prediction: that there should be fewer indeterminate responses to determinate relational problems (of the "three-term series" type) than to indeterminate ones. At first sight, this prediction seems trivial. Someone might be tempted to object that the reason why participants gave more determinate responses to determinate problems is simply that determinate problems have an obviously correct solution, whereas indeterminate problems do not. However, as we have pointed out, from the premises of indeterminate problems plenty of obvious, logically valid conclusions can be derived. It is just that they seem hardly worth deriving. So the objection might be revised: the reason why participants gave more determinate responses to determinate problems is simply that determinate problems have a solution that is both logically valid and worth deriving, whereas indeterminate problems do not. However, this revised objection misses the thrust of our argument. It presupposes but does not provide a solution to the problem we raised: When is a logically valid conclusion worth deriving?

It is only when some criterion for what counts as an acceptable solution has been explicitly characterised that the question of competence at reaching it can be raised. What we hope to have shown is that psychologists have been relying on an implicit criterion, based partly on explicit logical considerations, and partly on tacit pragmatic intuitions. Incidentally, these pragmatic intuitions have guided not only the work of psychologists of reasoning, but also, to some extent, the writing of logic textbooks—which in turn have influenced psychologists. Such textbooks waste little time in expounding the logical validity of derivations that are seen as too easy and too trivial to mention (except when they appear as steps in more "interesting" derivations). Typically, logic textbooks describe the conclusion of a deduction as being implicitly present or "hidden" in the premises —a description that ignores equally valid but trivial conclusions that are not hidden at all: for instance those that consist in the mere repetition of a premise, or the conjunction or disjunction of two premises. It is of course more interesting from a didactic point of view to teach students how to derive "hidden" conclusions, and, for the psychologist, to test people's ability at deriving them. However, it is important to recognise that this involves some reliance on an extra-logical, pragmatic criterion, which has itself to be spelled out.

Johnson-Laird and Byrne (1991, pp. 21–22) are exceptional in having addressed the issue of what conclusions seem worth deriving. They argue that, besides logical validity, there are three "extra-logical constraints" that govern the search for conclusions. Conclusions must not throw semantic information away, they must be parsimonious, and they must state "something new".

The first constraint, that conclusions must not throw semantic information away, rules out, for instance, inferring, from a premise of the form *P-and-Q*, a conclusion of the form *P-or-Q*, that would be less informative. The second constraint, that conclusions must be parsimonious, rules out, for instance, inferring, from a set of premises of the form {*If-P-then-Q*, *P*}, a conclusion of the form *P-and-Q*, when concluding just *Q* would be more parsimonious. The third constraint, that conclusions must state "something new" (glossed as "something not explicitly stated in the premises"), rules out, for instance, inferring, from a premise of the form *P-and-Q*, a conclusion of the form *Q-and-P* which would just restate the premise in another form. The goal pursued by Johnson-Laird and Byrne in formulating these constraints is intuitively clear and important. However, their formulation is not unproblematic.

The first constraint seems too strong, and is, in fact, in tension with the second constraint. When, in a *modus ponens* argument, one infers *Q* from a set of premises {*If-P-then-Q*, *P*}, as required by the parsimony constraint, one ends up with a conclusion that contains less semantic information than the premises, and which therefore violates the first constraint. More generally, deductive steps that, in a rule system, would be governed by elimination rules, generally involve some loss of semantic information. Deductions that do not involve any such eliminative steps, and which therefore respect the constraint of maintaining all semantic information, typically violate both the parsimony and the novelty constraints.

The parsimony and novelty constraints are intuitively more appealing than the constraint against loss of semantic information, but they are still rather vague. Consider for instance the premise *Joan has no brothers or sisters*. Suppose one were to derive from this premise the conclusion that *Joan has no siblings*. "Siblings" is shorter than "brothers or sisters", but is it psychologically more parsimonious? In terms of processing costs, there is the risk that the benefit linked to the shortness of "siblings" is more than offset by the cost of using a relatively rare word. Does the conclusion that *Joan has no siblings* express something new, i.e., something not explicitly stated in the premises? Intuitions may vary.

Another problem is that there may be many conclusions that satisfy the novelty constraint. For instance, from the premises {*Bill is better than Pete*, *Pete is better than John*}, the following two conclusions, among others, can be derived: *Bill is the best*, *John is the worst*. Both are novel in the intended sense. Are they equally likely to be derived? Is *John is the worst* (which involves an adjective not found in the premises) more novel than *Bill is the best*, and therefore more likely to be derived? The theory is mute on this point.

We share the intuitions underlying Johnson-Laird and Byrne's suggestion. What we have done is try to formulate these intuitions in a less problematic way by claiming that conclusions worth deriving are conclusions that are *relevant* in the precise sense of relevance theory, and the more relevant the better. In

particular, we argue, participants in a reasoning task assume that they are expected to come up with the most relevant conclusion available to them. The point of formulating things this way is not terminological, it is to help develop more precise, more explanatory, and more testable predictions. So what does make a conclusion worth deriving, and what counts as an acceptable response in a reasoning task? We have proposed the following general answer: an acceptable response is one that is relevant to the experimenter by demonstrating the participant's ability to expend the effort needed to derive a conclusion most relevant to themselves (under the pretence that they are reasoning from genuine, potentially useful information rather than from artificial experimental material). In the case of determinate problems, a response linking the two "end terms" of the premises obviously satisfies this criterion of acceptability. With indeterminate problems, there is no response that satisfies it in any obvious way. Nevertheless, some responses are more relevant than others and therefore more likely to be produced. This criterion allowed us not only to explain why determinate problems receive more determinate answers than indeterminate problems, but also to predict which determinate answers indeterminate problems would receive.

Manuscript received 30 March 2000
Revised manuscript received 25 October 2000

REFERENCES

Bernard, J.-M. (1998). Bayesian inference for categorized data. In H. Rouanet, J.-M. Bernard, M.-C. Bert, B. Lecoutre, M.-P. Lecoutre, & B. Le Roux, (Eds.), *New ways in statistical methodology* (pp. 159–228). Bern: Peter Lang.

Burt, C. (1919). The development of reasoning in school children. *Journal of Experimental Pedagogy*, *5*, 68–77.

Byrne, R.M.J., & Johnson-Laird, P.N. (1989). Spatial reasoning. *Journal of Memory and Language*, *28*, 564–575.

Carreiras, M., & Santamaría, C. (1997). Reasoning about relations: Spatial and non-spatial problems. *Thinking and Reasoning*, *3*, 191–208.

Clark, H.H. (1969a). Linguistic processes in deductive reasoning. *Psychological Review*, *76*, 387–404.

Clark, H.H. (1969b). Influence of language on solving three term series problems. *Journal of Experimental Psychology*, *82*, 505–514

De Soto, C.B., London, M., & Handel, S. (1965). Social reasoning and spatial paralogic. *Journal of Personality and Social Psychology*, *2*, 293–307.

Evans, J.St.B.T. (1995). Relevance & reasoning. In S.E. Newstead & J.St.B.T Evans (Eds.), *Perspectives on thinking and reasoning*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Evans, J.St.B.T., Newstead, S.E., & Byrne, R.M.J. (1993). *Human reasoning. The Psychology of deduction.* Hove, UK: Lawrence Erlbaum Associates Ltd.

Grice, P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Hayes-Roth, B., & Hayes-Roth, F. (1975). Plasticity in memorial networks. *Journal of Verbal Learning and Verbal Behavior*, *14*, 506–522.

Hilton, D.J. (1995). The social context of reasoning. *Psychological Bulletin*, *118*, 248–271.

Hunter, I.M.L. (1957). The solving of three term series problems. *British Journal of Psychology*, *48*, 286–298.

Huttenlocher, J. (1968). Constructing spatial images: A strategy in reasoning. *Psychological Review*, *75*, 550–560.

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

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

Moeser, S.D., & Tarrant, B.L. (1977). Learning a network of comparisons. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 643–659.

Mani, K., & Johnson-Laird, P.N. (1982). The mental representation of spatial descriptions. *Memory and Cognition*, *10*, 181–187.

Newstead, S.E., Manktelow, K.I., & Evans, J.St.B.T. (1982). The role of imagery in the representation of linear orderings. *Current Psychological Research*, *2*, 21–32.

Noveck, I.A. (2001). When children are more logical than adults: Investigations of scalar implicature. *Cognition*, *78*, 165–188.

Piaget, J (1921). Une forme verbale de la comparaison chez l'enfant. *Archives de Psychologie*, 141–172.

Politzer, G., & Macchi, L. (2000). Reasoning and pragmatics. *Mind and Society*, *1*, 73–93.

Politzer, G., & Noveck, I.A. (1991). Are conjunction rule violations the result of conversational rule violations? *Journal of Psycholinguistic Research*, *20*, 83–103.

Potts, G.R. (1972). Information processing used in the encoding of linear orderings. *Journal of Verbal Learning and Verbal Behavior*, *16*, 727–740.

Quinton, G., & Fellows, B.J. (1975). 'Percepual' strategies in the solving of three-term series problems. *British Journal of Psychology*, *66*, 69–78.

Roberts, M.J. (2000). Strategies in relational inference. *Thinking and Reasoning*, *6*, 1–26.

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, *57*, 31–95.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*, (2nd Edn.) Oxford: Blackwell.

Thompson, V.A. (2000). The task specific nature of domain-general reasoning. *Cognition*, *76*, 209–268.

Trabasso, T., Riley, C.A., & Wilson, E.G. (1975). The representation of linear order and spatial strategies in reasoning: A developmental study. In R. Falmagane (Ed.), *Psychological studies of logic and its development.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Van der Henst, J.B. (1999). The mental model theory and spatial reasoning re-examined: The role of relevance in premise order. *British Journal of Psychology*, *90*, 73–84.

Van der Henst, J.B. (2000). Mental model theory and pragmatics. *Behavioral and Brain Sciences*, *23*, 283–284.

Warner, S.A., & Griggs, R.A. (1980). Processing partially ordered information. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 741–753.