SCISPACE
formerly Typeset

Open access • Journal Article • DOI:10.1016/S0010-0277(00)00114-1

When children are more logical than adults: experimental investigations of scalar implicature. — Source link

Ira A. Noveck

Institutions: Centre national de la recherche scientifique

Published on: 01 Feb 2001 - Cognition (Elsevier)

Topics: Scalar implicature, Implicature, Pragmatics and Existential quantification

Related papers:

- Scalar implicatures: experiments at the semantics-pragmatics interface.
- Logic and conversation
- Some utterances are underinformative: The onset and time course of scalar inferences
- On the semantic properties of logical operators in English
- A Developmental Investigation of Processing Costs in Implicature Production

Share this paper:

View more about this paper here: https://typeset.io/papers/when-children-are-more-logical-than-adults-experimental-s3vf6dauyp

HAL
open science

When children are more logical than adults: Experimental investigations of scalar implicature

Ira Noveck

To cite this version:

Ira Noveck. When children are more logical than adults: Experimental investigations of scalar implicature. Cognition, Elsevier, 2001, 78, pp.165-188. hal-00000126

HAL Id: hal-00000126

https://hal.archives-ouvertes.fr/hal-00000126

Submitted on 20 Jan 2003

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.

Running Head: IMPLICATURE

When children are more logical than adults:

Experimental investigations of scalar implicature

Ira A. Noveck

Institut des Sciences Cognitives

Centre National de Recherche Scientifique

Lyon, France

Cognition (2001), 78/2,165-188.

Abstract

A conversational implicature is an inference that consists in attributing to a speaker an implicit meaning that goes beyond the explicit linguistic meaning of an utterance. This paper experimentally investigates scalar implicature, a paradigmatic case of implicature in which a speaker's use of a term like Some indicates that the speaker had reasons not to use a more informative one from the same scale, e.g. All; thus, Some implicates Not all. Pragmatic theorists like Grice would predict that a pragmatic interpretation is determined only after its explicit, logical meaning is incorporated (e.g. where Some means at least one). The present work aims to developmentally unpack this prediction by showing how younger, albeit competent, reasoners initially treat a relatively weak term logically before becoming aware of its pragmatic potential. Three experiments are presented. Experiment 1 presents a modal reasoning scenario offering an exhaustive set of conclusions; critical among these is participants' evaluation of a statement expressing Might be x when the context indicates that the stronger Must be x is true. The conversationally-infelicitous Might be x can be understood logically (e.g. as compatible with Must) or pragmatically (as exclusive to must). Results from five-, seven-, and nine-year-olds as well as adults revealed that a) seven-year-olds are the youngest to demonstrate modal competence overall and that; b) seven- and nine-year-olds treat the infelicitous Might logically significantly more often than adults do. Experiment 2 showed how training with the modal task can suspend the implicatures for adults. Experiment 3 provides converging evidence of the developmental pragmatic effect with the French existential quantifier Certains (Some). While linguistically-sophisticated children (eight- and ten-year-olds olds) typically treat Certains as compatible with Tous (All), adults are equivocal. These results, which are consistent with unanticipated findings in classic developmental papers, reveal a consistent ordering in which representations of weak scalar terms tend to be treated logically by young competent participants and more pragmatically by older ones. This work is also relevant to the treatment of scalar implicatures in the reasoning literature.

Keywords: Development, Implicature, Pragmatics, Reasoning

When children are more logical than adults:
Experimental investigations of scalar implicature

A participant in a laboratory reasoning experiment is comparable to a listener in a conversational exchange. Both are interlocutors who process incoming linguistic cues before eventually offering a reply. It is not surprising then that reasoning researchers often cite conversational, i.e. pragmatic, factors as sources of variability in their experiments. The best known of these factors is arguably conversational implicature (or Gricean implicature). This refers to an inference that consists in attributing to a speaker an implicit meaning that goes beyond the explicit linguistic meaning of an utterance. While the term implicature refers to a wide range of phenomena the present paper focuses on implicatures linked to critical logical terms (in this case, Might and Some). By focusing on a few terms in a structured context one can better determine the influence of implicature in general. This paper endeavors to show how implicatures can be investigated experimentally and to highlight what such inferences can tell us about cognition.

Although the notion of implicature may come up often in the reasoning literature, it is typically treated dismissively in one of two ways: a) as a kind of high-minded explanation for unanticipated responses or; b) as a phenomenon worth recognizing and then minimizing. For an example of (a), consider the following heading and paragraph from a paper on propositional inferences (Braine, O'Brien, Noveck, Samuels, Lea, Fisch & Yang, 1995).

> Nature of unpredicted responses
>
> ...A second major category of "other" responses can be interpreted as conversational implicatures or invited inferences. Under this heading we included inferences of the following forms: If not p then not q and If q then p from If p then q, inferences of not both p and q and If p then not q from p or q, and inferences of p or q and If not p then q from not both p and q. All of these were fairly common.

While the variety and extent of implicatures were duly cited, not much more about them was said (for another, similar treatment of implicature, see Johnson-Laird & Bara, 1984, page 23).

When implicatures have been directly investigated, it is to report that they are ultimately not relevant, thus falling into category (b). For example, Newstead (1995) reported that implicatures "seem to be virtually non-existent in syllogistic reasoning tasks." Newstead's negative claim about implicatures comes from valid syllogisms like the following:

All Nobel Prize winners are eminent scientists.

Some Nobel Prize winners are economists.

Some economists are eminent scientists.

That is, participants' not concluding that Some economists are not eminent scientists, which is arguably an implicature of the conclusion Some economists are eminent scientists, is taken as evidence for the non-existence of conversational influences in syllogistic reasoning. How Some are not can be inferred from Some and how this inference rears it head will be taken up in the paper. The point for now is that work like Newstead's (1995) leaves the impression that implicatures are not ubiquitous in laboratory tasks because participants do not mechanically produce conversational implicatures as errors in syllogistic reasoning tasks.

Not all cognitive scientists treat implicatures as phenomena of limited or questionable value. Linguists have been investigating implicature since it was introduced by Grice in the 1967 William James lectures and a small number of psychological studies have highlighted a role of implicatures in reasoning scenarios (e.g. Politzer and Noveck, 1991; Politzer, 1991; also see Hilton 1995). Although discussions concerning implicatures have evolved since Grice brought the notion on to the cognitive stage, it is worthwhile here to present his seminal proposal (see Grice, 1989).

The centerpiece of Grice's theory was the cooperative principle which exhorts participants in a conversation to:

> Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.

The principle itself does not give rise to particular implicatures, but he further proposed four maxims that do. Arguably, the best known of these is the quantity maxim (which comprises two submaxims):

i. Make your contribution as informative as is required.

ii. Do not make your contribution more informative than is required.

In Grice's original framework, he assumed that a hearer searches for implicatures when one of these submaxims is violated.[1] Consider the exchange in (1):

(1) Pierre: Are all the cakes ready?

Bettie: Some are.

Bettie's underinformative response violates quantity-submaxim (i) because she did not confirm his question with a "Yes" or with "All are ready." This compels Pierre to seek out an implicature along the lines of "Some of the cakes are not ready." Thus, Bettie did not say "Some are not," but most competent conversationalists would agree that this is derivable nonetheless (see Politzer, 1991, for experimental support).

This pragmatic analysis addresses the apparent paradox that arises when dealing with terms like Some. Whereas Some in such examples pragmatically prompts Some are not, Some is also logically compatible with All (Some can be glossed as Some and perhaps all). Thus, these two interpretations of Some conflict: Some are not is incompatible with perhaps all. Grice's initial proposal leads one out of the impasse by providing the framework in which one can distinguish between two compelling, but opposing, meanings -- one conversational and the other logical. That is, Grice can explain why such conversational inferences are accepted as valid, even if they are logically fallacious. This also highlights why this inference is relevant to reasoning research. In many reasoning studies, participants draw out Some are not from Some are and this is considered a logical fallacy.

Unsatisfied with Grice's explanation, neo-Griceans have analyzed this kind of implicature in some detail (e.g., Horn, 1973; Gazdar, 1979). According to these linguistic-pragmaticists, when one utters a relatively weak term (e.g. Some) , it is an indication that the speaker chose not to articulate a more informative term from the same scale (e.g. All). Presumably, the speaker does not know whether All is applicable or knows that it is not. Thus, uttering Some implicates Not All (and Not All is logically equivalent to Some are not). The scale of informativeness can be determined by entailment relations. The stronger term entails the weaker but not vice versa. All is a stronger quantifier than Some because All entails Some while Some does not entail All (to say that All Italians like ice-cream logically implies that Some Italians like ice-cream; however, Some Italians like ice-cream does not necessarily imply that All Italians like ice-cream). Given the prominent role of scales, this kind of inference has been dubbed scalar implicature and has since become a paradigmatic case for the study of implicature in the linguistic-pragmatic literature.

This neo-Gricean analysis is not restricted to quantifiers. It can be applied to a host of scales initially described by Horn (for a review, see Levinson, 1983). For example, if a speaker uses the modal Might (as in Bill might be in the office), it implies that the speaker had reason not to say the stronger-sounding Must (as in Bill must be in the office). Other scales can be applied to, but are by no means limited to, frequency (where the use of Sometimes excludes Always), epistemic status (where the weaker Think implies that it is not the case that Know), and another for connectives (where P OR Q indicates not both P AND Q). In each case, scales range from less to more informative and the speaker's use of a less-informative term implies the exclusion of a more-informative one.

The present study has two goals. One is to establish that scalar implicatures are psychologically real and common in reasoning scenarios. In addressing this goal, this work draws on theories of linguistic-pragmatics. To my knowledge there are no experimental studies that actually try to unpack the implicature process as described by contemporary linguistic-pragmaticists. It would be of obvious value to bring experimental data to bear on their analyses.

The second objective, which addresses the paper's experimental approach, is to establish how this class of weak scalar terms develop. Given the general thrust of Gricean and post-Gricean pragmatics, I will argue that the explicit meaning of a weak scalar term need be incorporated before any implicit meanings (i.e., implicatures) are determined. Thus, when a weak scalar term is employed (e.g. Some), its explicit meaning (e.g. at least one) must

be the default before it undergoes implicit pragmatic modifications (e.g. but not all). Given that the explicit meanings of these terms are tantamount to standard logic's, across development one should find indications of a consistent ordering in which logical meanings are preferred before implicit interpretations are.

How does one go about demonstrating the main developmental claim about implicature experimentally? I begin by pointing to evidence from the developmental literature that shows that children's initial representations of relatively weak scalable terms appear logical in nature before yielding to emergent pragmatic interpretations. This intriguing finding has been uncovered in three independent studies. Smith (1980) shows that younger children (four- to seven-year-olds), while appearing rather competent overall with quantifiers, "overwhelmingly" treat Some as compatible with All on a task where children had to answer questions like "Do some birds have wings"? This implies that the well-known pragmatic interpretation of Some (where Some is not compatible with All) arrives afterward. Similarly, with respect to propositional connectives, Braine and Rumain (1981) presented evidence showing that deductively competent 7- and 9-year-old children favor a logical interpretation of Or (which can be glossed as p or q and perhaps both) over an implicit one (p or q but not both). Adults on the same task were equivocal, though they tended to favor exclusive interpretations (Braine and Rumain, 1981, see Table 3). One other independent confirmation of this pragmatic effect comes from Paris (1973) whose data on disjunctions reveal the same developmental tendency as Braine and Rumain's. Thus, empirical findings indicate that interpretations of weak scalar terms among children, who are otherwise competent, are initially logical in nature and, with age, become potentially pragmatic. It should be noted that none of these authors anticipated such findings.

If this developmental tendency -- showing that pragmatic interpretations of weak scalar terms are consistently shown to increase subsequent to, and at the apparent expense of, logical meanings – can be generalized it would indicate the emergence of scalar implicature. It also leads to an unusual developmental curve in which young competent participants appear more logical than their older cohorts. This leads to the paper's experiments.

Experiment 1 tests participants' rendering of Might in a scenario that justifies two opposing treatments, a logical interpretation (where Might is compatible with Must) and a pragmatic one (where Might is not compatible with Must). The scenario also allows for the evaluation of an exhaustive set of modal statements in order to see how implicature development takes place with respect to logical development. Adults are anticipated to equivocate

between a logical and pragmatic reading of Might when a scenario allows for two possible interpretations. At around seven years of age, children have the tools to be equivocal as they become efficient at evaluating the import of two possibilities (see Acredolo & Horobin, 1987; Byrnes & Overton, 1986; Noveck, Ho & Sera, 1996; Sophian & Sommerville, 1988). Question is, Do these newly competent children behave equivocally in the face of two interpretations of Might? As pointed out earlier, prior literature indicates that the logical meaning of Might is likely to be the default among young, otherwise competent, participants and that the pragmatic meaning becomes available only with age. This implies that young competent reasoners will appear more logical than adults.

Experiment 2 is a follow-up to Experiment 1 and is designed to determine the extent to which the scalar implicature can be suspended. The training procedure of the modal reasoning scenario is modified slightly in order to favor logical interpretations. Implicatures were expected to be less apparent as cues to logical interpretations increase.

Experiment 3 investigates the relatively weak, existential quantifier Some (i.e., Certains in French). If the claim about implicature development is correct, then one ought to be able to generalize it to yet other weak scalar terms. Following up on Smith (1980), Experiment 3 investigates 8-year-olds, 10-year-olds and adults as it presents factually true or factually false statements with the quantifiers Some and All. A statement like Some giraffes have long necks is more likely to strike pragmatically-mature participants as wrong because they, more than their younger cohorts, would interpret the statement as Not all giraffes have long necks. The interest of Experiment 3 is to present converging evidence of the developmental pragmatic effect from a different sort of laboratory task.

Experiment 1: The modal Might

The paradigm in Experiment 1 is a reasoning scenario, but only as background for a puppet who utters an exhaustive series of modal statements. That is, modal statements expressing necessity, non-necessity, possibility, and impossibility are presented with respect to a visual scene. The scenario is described below.

Consider three boxes. One is open and has a toy parrot and a toy bear in it (the Parrot+Bear Box), the second is open and has only a parrot (the Parrot-only Box), and the third stays covered (Box C). Participants are told that Box C has the same content as either the Parrot+Bear Box or the Parrot-only Box. The paradigm may be explained in terms of a truth table. When a modal statement is true (or false) of both the Parrot+Bear box and the Parrot-only box, it is equally true (or false) of the Covered Box C. When a possibility statement is true of at least

one of the open boxes, it is true of the covered box. When a necessity statement is false of at least one of the open boxes it is false of the covered box (See Table 1).

Insert Table 1 about here

The puppet presents each of eight statements and it is the child's task to say whether the puppet's claim is right or not. The critical statement that allows us to study implicature is There might be a parrot in the box. On the one hand, if the participant adopts an explicit, logical interpretation of Might (where Might is compatible with has to), one would expect an affirmative reply ("the puppet is right"). On the other hand, if the participant adopts a pragmatic, restrictive interpretation for Might (where Might is not compatible with has to) one would expect a negative reply ("the puppet is wrong") or at least some equivocation. The structure of the task is identical to Noveck et al. (1996; Experiment 2), which tested for the influence of relative force and purposely by-passed the statement of interest here. Participants in this experiment were the same as those in Experiment 2 of Noveck et al.

Method

Participants

Thirty-two five-year-olds, 20 seven-year-olds, 16 nine-year-olds and 20 adult native English speakers participated. The children's mean ages (range) were 5;5 (5;1 - 5;11), 7;5 (7;1 - 8;0), and 9;4 (9;0 -9;5). Participants were recruited from Minneapolis and St. Paul in Minnesota. Adults participated to fulfill requirements for the Introductory Psychology course at the University of Minnesota.

Materials

Two open boxes and one closed box were presented. One open box contained a parrot and another contained a parrot and a bear. Participants were told that the closed box had the same contents as one of the two closed boxes (see Procedure). Participants then heard eight statements by a puppet named Wylbur: (1) There has to be a parrot in the box (true), (2) There does not have to be a parrot in the box (false), (3) There might be a parrot in the box (true), (4)There cannot be a parrot in the box (false), (5) There has to be a bear in the box (false), (6) There might be a bear in the box (true), (7) There does not have to be a bear in the box (true), (8)There cannot be a bear in the box (false). Four random orders of the statements were prepared.

Design and Procedure

Training began by acquainting the participants with the puppet and by stating that the puppet says things that are right sometimes and wrong sometimes. To provide an example, Wylbur would say that "the wall is red" in reference to the white Experimental room. Children were then asked, "Is Wylbur right?" The same technique was used to provide a situation in which Wylbur was indeed right.

Participants were then shown four boxes. Two boxes were open: one box contained a horse and a fish (Horse+Fish Box) and the other contained just a horse (Horse-only Box). The two other boxes were covered but were identical to the Horse+Fish Box and the Horse-only box. Participants were shown the two open boxes and asked to tell the experimenter what they saw. Participants were then shown one of the covered boxes (chosen at random) and heard the Experimenter say, "A friend of mine gave me this box and said 'all I know is that whatever is inside this box looks like what's inside this Box (Experimenter pointed to the Horse+Fish Box) or what's inside this Box (Experimenter pointed to the Horse-only Box).'" Participants were then instructed to peek inside the box, but to not say what they saw (they saw either a horse alone or a horse and a fish).

At this point, the puppet would say "There is a fish inside the box." The child was asked to say whether the puppet was right. If the child had peeked in the Horse+Fish Box, she would be expected to say that the puppet was right; if the child had peeked inside the Horse-only Box, she would be expected to say that the puppet was wrong. The same procedure was carried out on the other, closed box: the child peeked inside and the puppet said "There is a fish inside the box." Errors, which were very rare here, were pointed out only after both of the initial two questions were presented.

Before the third and last statement of the training session was presented, a child was asked to close her eyes as the Experimenter placed, at random, one of the closed boxes in front of the participant. The participant was told that to reply next time, she did not need to open the box. The participant then heard the puppet say "There is a horse inside the box." Typically, children answered correctly by agreeing with the puppet. If there were errors at this point in the training, the procedure was carried out again. Seven-year-old and older participants had no difficulty at all with the training. Five-year-olds occassionally required more than one turn with this last training statement (such a participant would indicate that the only way to answer was to look inside the covered box before making a judgement).

Experimental trials concerned larger and differently colored open boxes. One box contained a Parrot and a Bear (hereafter referred to as the Parrot+Bear Box) and the other contained just a Parrot (the Parrot-only Box). As in the training session, children were shown a box that was covered and were told that "a friend of mine gave me this box and said 'all I know is that whatever is inside this box looks like this Box (Experimenter pointed to the Parrot+Bear Box) or this Box (Experimenter pointed to the Parrot-only Box).'" However, unlike in the training session, this was the only covered box children encountered and they were told that they would not be able to look inside of it until the "game" was over.

Results

Analyses begin with an overview of children's reasoning abilities on this task. This way one can determine the approximate age at which children show competence before I turn to the expression of interest, There might be a parrot in the box.

Table 2 shows participants' rates of correct evaluations to each of the eight modals statements across the four age groups. Participants' performance is of interest with respect to: 1) chance --probability is .5 that a child would be correct on any given statement and; 2) developmental changes. The table shows that five-year-olds are above chance levels in three of the eight conditions. This is an impressive rate of success, but it does not reveal that the five year olds have largely mastered the task. Seven-year-olds answer correctly at rates that are above chance levels for seven out of the eight conditions and show the earliest signs of showing consistent mastery on the task. The same holds for nine-year-olds. That seven-year-olds should appear competent in modal reasoning abilities conforms to expectations based on the literature (for a review, see Braine and Rumain, 1983).

Insert Table 2 about here

Seven-year-olds' rate of logical interpretations with respect to There might be a parrot in the box (80%) is intriguing not only because they respond at rates that are significantly above chance levels but because they do so at a rate that is significantly higher than that of the adults' (35%), $\chi^2 = 8.28$, $p < .01$. Most adults assume that the possibility that the parrot will be found in the hidden box is wrong because the expectation is that the parrot's presence in the hidden box is necessary. Nine-year-olds look less like the seven-year-olds; 69% provide the

logically correct answer. Nevertheless, the difference between nine-year-olds and adults is significant, $\chi^2 = 4.15$, $p < .05$.

Note that the seven-and nine-year-olds are doing some rather sophisticated modal reasoning and they have no difficulty rejecting statements that appear wrong. A significant number of children detect logically wrong statements like there cannot be a bear and there does not have to be a parrot. Yet, these same children tend to agree with there might be a parrot. Their ability to detect wrong statements does not extend to those that would involve pragmatic interpretations. Adults also detect when statements are wrong but their equivocality surrounding there might be a parrot indicates that they are wary of its potential for two interpretations.

For those participants whose performance was otherwise perfect (for the remaining seven statements), the ratios between the logical and pragmatic readings of might are especially illuminating. Of six otherwise error-free seven-year-olds, four reveal using a logical interpretation and two a pragmatic one (2:1 ratio). Of ten otherwise error-free nine-year-olds, six reveal a logical interpretation and four a pragmatic one (3:2 ratio). In contrast, of nineteen otherwise error-free adults, six reveal a logical interpretation and 13 a pragmatic one (a ratio that is roughly 1:2). The seven- and nine-year-olds generally treat might as compatible with has to and adults tend to treat might as not compatible with has to. (Furthermore, there was one precocious, error-free five-year-old who treated might logically). The most plausible conclusion concerning the developmental effect appears then to be that implicature is, at least in the present scenario, not readily available until children are older. The response to the There might be a parrot in the box statement in this context provides the cognitive literature with an unusual developmental curve.

Discussion

Seven-year-olds tend to accept the logical interpretation of There might be a parrot in the box whereas adults tend to cash out its pragmatic potential. This finding can be taken to mean 1) that the growing conversationalist effortfully prefers a logical interpretation over a pragmatic one or; 2) that logical interpretations of might remain the default and that they give way to pragmatic interpretations. I consider the second explanation more plausible. Although young children's reasoning and discourse skills have long been documented, it is another matter to suppose that seven-year-olds -- in deciding between a logical and pragmatic inference -- opt effortfully for the logical interpretation while adults do not. It seems more probable that the pragmatic

interpretation is simply less accessible to the younger participants than it is for the older ones. One would expect children to do less work than adults and the children's answers indicate how far they go in their reflections. This conforms to the findings regarding the existential quantifiers and disjunctions cited earlier.

In presenting these results, I have heard other interpretations that try to account for the seven- and nine-year-olds' tendencies with respect to the statement There might be a parrot in the box. I would like to describe two of these and show why they are ultimately unsatisfactory. The first echoes a possibility that I have considered myself: Perhaps younger children are just looking for a quick verification. They hear Might and determine quickly that the given statement can be verified regardless of the genuine possibility or necessity of the statement. The presence of a parrot somewhere among the options verifies the statement. There are two difficulties with this proposal. First, even if this characterisation of children's reactions were accurate, this is just another way of saying (albeit in processing language) that children entertain the logical interpretation of Might and go no further: They recognize that Might meets a minimum standard and, satisfied, they do no further processing. Secondly, the characterisation of children in this experiment as impatient satisficers is not supported by the data. There is no indication that children are reacting rashly or impatiently to the statements generally. The 7- and 9-year-olds demonstrate some rather sophisticated, thorough-going reasoning skills to all the other, non-equivocal statements. How else could these children respond correctly to statements like There might be a bear, There does not have to be bear, There has to be a bear, There cannot be a parrot or There cannot be a bear unless they were doing some exhaustive work and considering all possible outcomes of this scenario? For example, one would expect chance responding to There might be a bear among seven-year-olds if they were quickly verifying just one of the boxes.

A second possible account argues that implicature is equally available for children and adults, but that the context more readily cancels it for children than for adults. This account accepts the scalar implicature explanation, but assumes that it is more prominent than claimed here, even among children. If this were so, however, it is not clear why the context is sufficient for cancelling the implicature for children and not for adults. Cancelling is an extra step that one would more likely attribute to more sophisticated participants and not the other way around.

The scalar implicature in the present study can arrive only under particular circumstances, i.e. when an utterance includes a weak scalar term that is underinformative relative to the context. This explains why the

statement There might be a parrot in the box is of interest to this Experiment and why hardly any of the other items concerning the parrot lead to unusual developmental patterns. The late-developing scalar implicature account is both theoretically driven and is supported by similar findings in the literature. It is more plausible and parsimonious to simply assume that interlocutors accept an utterance as true unless they have reasons to object. The otherwise-competent seven- and nine-year-olds generally do not see what is objectionable in the statement There might be a parrot in the box whereas adults do. Specifically, it is proposed that in hearing There might be a parrot, adults more readily than children infer the scalar implicature.

Experiment 2: Favoring logical interpretations

Experiment 2 aims to verify the main findings of Experiment 1. However, it presents the experiment after providing more intensive training. It was hypothesized that a more thorough understanding of the task, led by queries focusing on individual items in the hidden box, would encourage logical interpretations. After successful completion of the training, participants were presented the same scenario as in Experiment 1 and, this time, with two groups of modal terms. If the conclusions reported after Experiment 1 are confirmed, one should find that 5-year-olds appear less competent than the seven-year olds, that seven-year-olds demonstrate adult-like sophistication, but that adults still reveal evidence of producing an implicature while the seven-year-olds do not.

Method

Participants. Nineteen five-year-olds (mean age: 5;5, range: 5;1 - 5;11), sixteen seven-year-olds (mean: 7;7, range: 7;5 - 7;11) as well as sixteen adult native English speakers participated. The children were either recruited or came from a private school in the Twin Cities area. Adults were compensated monetarily to participate.

Design and Procedure. Part I. The training task was carried out in a manner identical to that in Experiment 1. That is, one exposed box had a horse alone and another had a horse and a fish. After the training was completed, the Experimenter put the puppet down and said "Now, I would like to ask you a few questions about the covered box." These are listed and numbered below (though two random orders of queries were prepared):

1. Could the box be empty? (Answer: no)
2. Could there be a cat in the box? (no)

3. Could there be a fish by itself in the box? (no)
4. Could there be a horse by itself in the box? (yes)
5. Could there be a horse and a fish in the box? (yes)
6. If we open the box, could there be a horse inside? (yes)
7. If we open the box, could there be a fish inside? (yes)

The first question was included to verify that children do not consider the null possibility simply because the question compels a participant to consider it. The second question addresses the same concerns but also verifies that participants understand that there must be something in the box. The third question is presented to verify that participants do not make the error that 15% of adults made in a prior follow-up [2]. The fourth and fifth questions are included to verify that children understand the parameters of the task as intended. The last two questions are presented to verify that children are able to consider a box's item independently even if it may appear with the other item.

If participants erred on any one question, the error was pointed out at the end of the round and the entire set was readministered. Participants had three opportunities to answer all the questions in a round. Otherwise, their responses were not included in the analyses (which occurred among three five-year-olds).

Part II. Evaluations of the modal statements were presented as they were in Experiment 1. The only difference is that two sets of statements were presented. Each set was presented by one of two puppets. Each set was prepared so that one can employ modal terms and their negatives. For example, <u>might</u> and <u>might not</u> (a modal expression that was not included in Experiment 1) were included in one set and <u>has to</u> and <u>does not have to</u> were included in the other. One set evaluated children's comprehension with the following modal expressions:

<u>There has to be a</u> (Necessity)

<u>There does not have to be a</u> (Non-necessity)

<u>There could be a</u> (Possibility)

<u>There could not be a</u> (Impossibility)

The other set evaluates the parrot and the bear's presence with:

<u>There must be a</u> (Necessity)

<u>There might not be a</u> (Non-necessity)

There might be a (Possibility)

There must not be a (Impossibility)

Results and Discussion.

Five-year-olds were the only ones to reveal having any difficulty with the training. On average, they needed to hear the set of questions 1.94 times (excluding three participants who were dropped) compared to 1.06 and 1.00 times for the seven-year-olds and adults, respectively. The most difficult question If we open the box, could there be a fish inside? yielded No responses among 50% of the five-year-olds. When asked about the "No" response, these participants were wont to say that there had to be a fish and a horse. If this question were removed from the analysis, 81% of the 16 competent five-year-olds' responses would have succeeded on the training after just one round. The excluded three had general difficulties.

Insert Table 3 about here

Table 3 presents participants' rates of correct evaluations of the modal statements. These are first analyzed to test for internal consistency. This is followed by analyses to verify the findings of Experiment 1.

The two sets of synonymous modal statements yield rates of correct responses that are remarkably alike. In all but one of twenty-four comparable cells, rates of correct responses to one modal statement is similar to its equivalent modal statement at each age level and with regard to each animal. For example, the statement There has to be a parrot in the box yields rates of correct responses similar to There must be a parrot in the box at each age level and There could not be a parrot in the box yields rates of correct responses similar to There must not be a parrot in the box at each age level. The only exception is adults' responses to There could not be a bear in the box (50%) vs. There must not be a bear in the box (94%), $t(15)= 3.41$, $p < .01$. The word could is the source of some variation among adults. The role of could in this experiment will be taken up in the General Discusion. Overall, the percentage of times five-year-olds, seven-year-olds and adults give the same response to both of the equivalent statements is 69%, 91%, and 90%, respectively. Here, again, one can see how there is a behavioral gap between five- and seven-year-olds.

To what extent does intensive training affect performance with respect to the findings of Experiment 1? Rates of correct responses from Experiment 1 were compared to those reported here with has to, does not have to and might (modal expressions that were employed in Experiment 1) and must not. A 2 (Training: Basic vs. Intensive) X 3 (Age: five-year-olds, seven-year-olds, and adults) ANOVA was computed with the total number of correct responses treated as the dependent measure. There is a main effect for Training, $F(1,118) = 4.982$, $p < .05$ and, not surprisingly, there is a main effect for Age, $F(2,116) = 54.396$, $p < .001$. Thus, the extra attention paid to the training task increased rates of logical responding.

The intensive querying reduced the difference between the seven-year-olds and adults with respect to the statement There might be a parrot in the box. The extra training appears to have encouraged logical responses among adults and at a rate that is significantly greater than that in Experiment 1, $\chi^2 = 5.84$, $p < .05$. Also, the logical:pragmatic ratio (a logical vs. pragmatic reading for those who gave correct responses otherwise) was high among the adults (11:2) and it was similar to the seven-year-olds' (6:1). Even though the adults' responses are encouraged to become logical (75%), it is still not altogether at ceiling and the rate of logical responding is still higher (albeit non-significantly) among the seven-year-olds (94%).

Overall, Experiment 2 shows that the age trends found earlier remain stable. Five-year-olds' low rates of success remains and by the time children are seven years old, they appear competent in making modal judgments. The intensive training increased the likelihood that adults would provide a logical response to the statement There might be a parrot in the box. Shades of the developmental effect associated with this statement, as reported in Experiment 1, remain evident.

Experiment 3: The French existential quantifier Certains

As the Introduction noted, other studies point to a developmental effect similar to the one in Experiment 1. For example, Braine and Rumain reported that younger children have a tendency to treat the disjunction or inclusively (where or is compatible with the conjunction, as in p or q and perhaps both) while adults have the tendency to treat the disjunction exclusively (as in p or q but not both); similarly, Smith reported that four- to seven-year-olds treat Some as compatible with all, but she did not produce adult data. The implication of Smith's work is that older children and adults would eventually note that Some pragmatically implies not all.

Experiment 3 continues where Smith left off. To demonstrate that the reported pragmatic effect is robust, Experiment 3 employs Smith's original paradigm (with two minor modifications -- it removes the question form and it is presented in French). Smith's experiment brings factual knowledge into play as it presents participants with a series of sentences having either the form Some X [verb] Y or all X [verb] Y. Her experiment was designed so that the content for many Some sentences (e.g. those like Some elephants have trunks) would still be true even if the quantifier were all; other Some sentences (e.g. Some birds live in cages) would not be true if the quantifier were all. She reported that Some was compatible with all for her young participants because they responded in the affirmative to (infelicitous) statements like Some elephants have trunks. From a pragmatic point of view, it appears that her young participants did not draw the implicature.

The prediction for the present experiment is the following: Given that the pragmatic interpretation of Some (not all) is proposed to arrive subsequent to the logical interpretation, the infelicitous Some sentences are predicted to yield more negative ("no, I disagree") responses as participants become older.

This Experiment, like Smith's, includes two controls with respect to the materials. Companion sets (A and B) were prepared so that each kind of content (e.g., elephants having trunks) can be presented with the quantifier Some or the quantifier all (and an individual participant heard one or the other). This is to verify that, overall, participants draw the same conclusions about which universally and existentially quantified propositions are indeed true, regardless of content. If a participant received set A with universal quantifiers then she would receive set B with existential quantifiers and vice versa. The other control is that each participant heard (the same set of) absurd factual statements (e.g., all chairs tell time/Some garage doors sing).

One new control introduced here is that the study was conducted in a double-blind manner. The experimenter was instructed to read each statement in the same fashion throughout and without undue emphasis on any one word. The author did not reveal the purpose of the study until all the data were collected (and the absurd control statements fortuitously provided adequate cover). This method avoids a possible criticism concerning the potential of participants to detect prosodic cues or the experimenter's intention.

Method

Participants

Thirty-one eight-year-olds, 30 ten-year-olds, and 15 adult native French speakers participated. The children's mean ages (range) were 8 years and 2 months (7;4 - 8;11), and 10;7 (10;0 - 11;7), respectively. Child participants were recruited from each of two grades in schools located in Domancy and Sallanches, two small villages in the Haute Savoie region of France (in the French Alps). Adults were volunteers from the same villages as the children.

Materials

Smith's (1980) experiment served as a model. In fact, a substantial portion of the present materials was incorporated directly from her Appendix. Sentences were based on three types of information: factually universal (that elephants have trunks is arguably best represented with the quantifier all), factually existential (that birds live in cages is arguably best represented with Some), and absurd (that stores are made of bubbles is arguably false with both kinds of quantifiers). The materials were made up of 30 sentences describing a relation; these can be broken down into the following subgroups:

a) 5 absurd all sentences (e.g. All chairs tell time)

b) 5 true all sentences (e.g. All elephants have trunks)

c) 5 false all sentences (e.g. All dogs have spots)

d) 5 absurd Some sentences (e.g., Some stores are made of bubbles)

e) 5 true (and felicitous) Some sentences (e.g. Some birds live in cages)

f) 5 true (but pragmatically infelicitous) Some sentences (e.g. Some giraffes have long necks)

The ten sentences in (a) plus (d) remained stable across participants. In contrast, one can see how a switch in quantifiers can make (b) interchangeable with (f) as well as (c) with (e). All 15 Some sentences were presented as a series as were the 15 all sentences (order of Some vs. all was varied). Otherwise, two random orders for each group of 15 Some and 15 all sentences were prepared. This results in four sets of materials. The materials in one of its forms is presented in English (translated from French) in the Appendix. The materials were designed so that two sentences in groups (b) and (e) concerned natural categories and that three concerned artifactual categories. In this way the resulting switched sets remain comparable in terms of content (in fact, subsequent analyses of the results revealed that these factors made no difference).

Procedure

Participants were told that they were going to be presented a series of statements and that it was their job to simply say whether or not they agree with each. They were told that it was not a test and that they would occasionally be instructed to explain their response. Two examples were presented that were based on visual cues: Some of the pens are red (which was true) and all of my buttons are black (True). It was anticipated that participants might react to the absurd statements with incredulity. Thus, the experimenter was instructed to simply remind participants to respond with yes, I agree or no, I do not agree whenever the participant appeared unsure. The experimenter reported no particular difficulties with the materials.

Results

The results from this experiment are presented in Table 4. If one excludes only adults' performance with the pragmatically infelicitous statements, rates of correct responses are above levels predicted by chance in every cell of the Table. Clearly, the 8-year-olds and the 10-year-olds answer consistently over the quantifiers certains (Some) and tous (all). It is important to point out that children show themselves to be highly competent with respect to the remaining statements. This provides the context within which one can evaluate the performance on the pragmatically infelicitous Some statements. Of course, this is not surprising given that Smith (1980) reported sophisticated responding among four- to seven-year-olds on the same kind of task.

Insert Table 4 about here

In order to analyze the statement of interest (the last row of the Table), a 3 (Age: 8-year-olds, 10-year-olds and adults) X 2 (Sets of Statements) X 2 (Presentation order of quantifier) ANOVA was carried out with percentage of logically correct responses serving as a dependent measure. The results were very neat, showing one main effect for age, $F(2,64)=11.898$, $p< .0001$. There were no other significant effects. Post-hoc tests show that the significant difference is due to the rates of performance that differ between the 10-year-olds and the adults.

Adults' performance generally revealed that they either disagree with all of these statements (six participants, 40%), agree with one or two of the statements (four participants, 27%), or treat the statements as true by agreeing with all five of these statements (five participants out of fifteen, 33%). In contrast, only 2 of 30 ten-year-olds (6%) disagreed with all of the statements, two (6%) agreed with just one of them, five (17%) agreed with

four of the items and 21 agreed with all five statements (70%). The eight-year-olds' response patterns showed that they were even more likely than the ten-year-olds to accept the logical interpretation of the Some statement. Only one participant appeared pragmatically aware by disagreeing with all five of the infelicitous statements (3%), eight (26%) agreed with four of five statements and twenty-one (68%) agreed with all five pragmatically infelicitous statements.

Discussion

This task is relatively difficult. Participants were required to evaluate quantified statements while drawing on working memory. However, one finds children operating rather competently and in line with the prediction: pragmatic interpretations become evident subsequent to logical interpretations.

As in Experiment 1, the children's responses indicate that they are thorough and consistent over the course of the experiment. Not surprisingly, children correctly rejected the absurd Some statements (like Some stores are made of bubbles), indicating that they were not apt to fantasize. Note too that children were very good at detecting the appropriateness of quantified sentences when the statements were not infelicitous. Especially interesting are statements like Some birds live in cages, which were correctly accepted, and statements like All birds live in cages, which were correctly rejected; thus children are able to detect when the factual aspects of a statement are compatible with Some but not all. However, the same children did not reject infelicitous statements like Some giraffes have long necks. Such rejections are common only among adults. It appears that children are being rather thorough in all their judgements and, yet, pragmatic interpretations are generally not evident to them on these tasks.

General Discussion

Grice's account of implicatures, as being governed by the Co-operative principle and the associated maxims, has been either accepted or refined, but rarely rejected. The experiments presented here aimed to establish the anticipated psychological reality of Gricean implicatures by investigating a paradigmatic case -- scalar implicature -- largely from a developmental perspective. The experiments succeeded in demonstrating, not only that these Gricean implicatures are present in adult inference-making but, that, in cognitive development, they occur only after logical interpretations have been well established. One finds that linguistically and inferentially competent seven- and nine-year-olds do not readily endorse a pragmatic interpretation of the term Might whereas

adults do. The same holds for 8- and 10-year-olds with respect to Some (Certains in French). Below I summarize the main findings, consider their implications, and determine how well Grice's theory anticipates them. This is followed by analyses from two post-Gricean theories and their accounts of the experiments' main results.

Three features of these studies are noteworthy. One, this study's most intriguing result is not entirely unexpected. Indications from the literature were that children tend to not attach an upper bound when treating weak scalar terms like Some and Or; they tend to treat Some initially as compatible with All and allow for an inclusive-or interpretation of the disjunction. Two, the children are competent in all other aspects of the two tasks. Thus, it cannot be argued that children are randomly choosing one interpretion over another. Nor can it be argued that children are confusing Might with Has to or Some with All. Children are rather strong at recognizing when utterances are overinformative relative to the given situation or fact (as in All birds live in cages) but are not so determined when they are required to detect an infelicitous underinformative utterance (as Some elephants have trunks). This speaks directly to the issue of scalar implicature, which is normally discussed in the context of underinformativeness. The prerequisites for scalar implicature are present, but the young participants' responses suggest that these conditions are not being taken into account. Three, note that the scalar implicature, or lack thereof, occurs within a well-defined context, i.e. where there is derivable background information that justifies a stronger characterisation. If Experiment 1 presented There might be a parrot in the box in a decontextualized setting, it is not clear what participants would do. Future experiments could determine the situations in which implicatures become evident.

Although particular ages are noted -- e.g., that seven-year-olds were the youngest children to demonstrate modal competence while not appearing to make pragmatic interpretations of Might -- I have carefully avoided making the claim that the general capacity to infer implicatures is linked to a particular chronological age. The paper is only claiming that the competent use of a weak scalable term is linked initially to an explicit interpretation and that this is followed by a pragmatic one. This is drawn out by the results. By no means do I want to suggest that children are incapable of pragmatic inferencing at younger ages. The tasks described here, which are typical of those found in the developmental literature, require no small amount of work as they require children to compare an utterance to real-world knowledge. This might well mask an ability to do pragmatic inferencing at younger ages. In fact, children younger than seven carry out other sorts of pragmatic inferences (Shipley, 1979; Stedmon,

1986; see Politzer, 1986, for a thorough discussion). I argue that one would find the same effect among even younger children if a task were made easy enough. However, this latter claim goes beyond the scope of the paper.

Given that Grice was the originator of many of the concepts studied here, it is appropriate to determine how well his insights are captured in these settings. As described in the Introduction, Grice would probably have argued that adults say that a statement like There might be a parrot in the box is "wrong" because they are alert to the first submaxim of quantity. Adults are objecting to the assertion of Might because a more informative term (for necessity) is contextually justified; i.e., it is inappropriate to use a less informative term. The developmental effect in Gricean terms would be that adults are more likely to detect a violation of the quantity maxim than children are. Once the violation is detected some kind of implicature is produced that reconciles the utterance with the cooperative principle. The exact content in this scenario is not made clear by Grice's theory, but the violation is assumed to be a signal to the listener that the speaker is conveying more than the explicit meaning of the utterance.

Two properties of Gricean implicature are relevant in the present work. One is that scalar implicatures, once available, are not obligatory. In Gricean terms, the implicature is defeasible. This property can be cited in explaining how the percentage of logical responses increased among adults (significantly in Experiment 2) as the task requirements were modified to focus attention on the range of possibilities in the modal task. [3]

The second Gricean property that is pertinent to the present work is non-detachability. This is a (complicated) means of saying that implicatures, although not part of the lexical meanings of terms, are normally carried, in the same context, by synonymous expressions. Thus, if one were to replace Some with a subset of, one would still expect a related implicature (something like subset and not the entire set). Experiment 2 presented a situation that failed to support non-detachability. The modal could did not appear to prompt an implicature similar to the one for Might. This implies either that could is not a proper synonym of Might or that Grice's non-detachability condition is too strong. Of these two possibilities, the first is more likely because it is difficult to imagine what would be the term that a) both shares a scale with could and is on par with necessity, and that b) is denied in an implicature. Others have argued that the non-detachability condition is not plausible (Levinson, 1983). Nevertheless, Grice's construct has largely been validated; its weakness is that it does not predict the specific kind of implicature produced.[4]

I now turn to two modern pragmatic accounts and the way each would address scalar implicature in particular. I refer to one as the neo-Gricean account while the other is Relevance Theory. Each would have to explain two findings. One is the main developmental finding in the Experiments and the other is the way evidence of implicature is reduced in Experiment 2.

The neo-Griceans (e.g., Horn, Gazdar, and Levinson) are credited with developing theoretical accounts of scalar implicatures that provide more precision than Grice's. Part of their description assumes that scalar implicatures are produced rather mechanically on weak scalar terms; in other words, these implicatures are assumed to be generalized (as opposed to being situation-specific or particularized).[5] The neo-Griceans would probably argue that the developmental ordering, which indicates that scalars are derived and not part of the logical meaning, shows that the implicature is automatized with time. Thus, younger participants who do not reveal having made the implicature have not derived it with regularity yet. With regard to adult uses, a generalized neo-Gricean account need explain those instances where scalar implicatures are not in evidence. This is discussed often in the linguistic literature (e.g., Cherchia, 1999). These theorists would argue, in line with Grice, that implicatures are often blocked or rendered defeasible. The manipulations of Experiment 2 have blocked the implicature for the adults.

A second proposal for scalar implicatures comes from Carston's (1998) application of Relevance theory (Sperber & Wilson, 1986/1995), an alternative pragmatic theory that actually began as an attempt to improve on Grice's. In the Relevance framework, an implicature is defined as an inference that the speaker intends and expects the hearer to draw in order to arrive at an interpretation of the utterance that is relevant enough. In particular, a scalar implicature is derived when a relatively weak statement fails to meet the hearer's expectation of relevance. For instance, in the dialogue:

(1) Pierre: Are all the cakes ready?

Bettie: Some are.

Bettie's answer is not relevant enough unless it is taken to implicate that some of the cakes are not ready. According to neo-Griceans, scalar implicatures are automatically derived by competent language users, and can then be cancelled if the context suggests doing so, whereas, for Relevance theory, scalar implicatures are derived only when they are contextually needed to achieve the expected level of relevance (which is

quite commonly the case). Relevance theory might then propose that the competent children who failed to derive a scalar implicature had their expectations of relevance satisfied with the logical meanings of Might and Some. This might be because children's expectations of relevance, at least in the context of these experiments, are more easily satisfied than that of adults. Or it might be that the cost of deriving these implicatures is greater for children, thus offsetting their cognitive benefit and thereby lowering their relevance. The fact that, according to the context, adults opt for a logical or a pragmatic interpretation of Might and Some is just what Relevance theory predicts and calls for no ad hoc explanation in this framework.

The present experiments were designed to capture scalar implicature in experimental settings in order to identify and characterize critical features of a phenomenon that is of central importance to the linguistic-pragmatic literature. They were not designed to determine which theoretical account (in linguistic-pragmatics and beyond) better anticipates the main results reported here. Our evidence is compatible with either the neo-Gricean or the Relevance account of scalar implicatures and it would be a considerable advance if one were to design further experiments to test between the two accounts. For the short term, I hope to have established empirical facts upon which implicatures can be investigated further, no matter one's theoretical orientation.

From a psycholinguistic perspective, Politzer's (1986) "Conflict hypothesis" was the first to point to the tension between (standard) logical responses on reasoning tasks and the potential for pragmatic readings. Through a review of both the developmental and adult literature, he showed that many so-called errors on reasoning tasks are reasonable responses in light of linguistic-pragmatic considerations. The present work furthers this approach by isolating scalar implicatures and uncovering their potential for influence on typical laboratory tasks.

It has been over 30 years since Grice brought his seminal notion of implicature on to the cognitive stage and it has inspired much work since. Linguists have isolated paradigmatic cases to analyze and cognitive psychologists have (perhaps unwittingly) justified some of their claims. The present study has attempted to engage these two, largely independent, strands of research. It points to the psychological reality of scalar implicatures and shows how a linguistic-pragmatic analysis can account for an intriguing developmental phenomenon.

Acknowledgements

This work was supported by grants from the (U.S.) National Institute of Child Health and Human Development (T32 HD-07151) while the author was at the University of Minnesota as well as by a grant from the CNRS (ATIPE, 1999). Versions of this paper have been presented at the Linguistics Association of Great Britain (Luton, 1998), the European Society of Philosophy and Psychology (Warwick, 1999) and the International Symposium on Linguistics (Lyon, 1999). I wish to express my gratitude to Dan Sperber, especially, for his continual support and his multiple suggestions as well as to the other members of le GRICE (Groupe de Recherche sur l'Inférence et Compréhension Elementaire) for their ready feedback: Jean Baratgin, Luca Bonatti, Guy Politzer, and Jean-Baptiste van der Henst. I also wish to thank the following people whose comments or encouragement were crucial: Deirdre Wilson, Larry Horn, Mira Ariel, Jacques Moeschler, Pierre Jacob, Manuel Garcia-Carpintero, Gennaro Chierchia, Maria Sera, and the late Martin Braine. Finally, merci à Corinne Gruel for carrying out Experiment 3.

Correspondence concerning this article should be addressed to Ira Noveck, Institut des Sciences Cognitives, 67 Blvd. Pinel, 69675 Bron, France (email: noveck@isc.cnrs.fr).

Notes.

[1] The other maxims are:

Maxim of Quality

1. Do not say what you believe to be false.
2. Do not say that for which you lack adequate evidence.

Maxim of Relation

Be relevant.

Maxim of Manner

1. Avoid obscurity of expression.
2. Avoid ambiguity.
3. Be brief.
4. Be orderly.

[2] When twenty adults were presented the same task as Experiment 1's, but with the response options true or false, 15% incorrectly said that "There has to be a parrot in the box" was false and that "There does not have to be a parrot" was true. Protocols revealed that 3 of 20 adult participants rejected a premise and assumed that it was possible to have a bear in the box by itself.

[3] Newstead's results, as reported in the Introduction, are not inconsistent with Gricean claims and for the following three reasons. First, it is not clear why one would expect participants, who in his experiments were required to provide (or evaluate) one response per syllogism, to replace anticipated logical deductions with conversational implicatures. This expectation implies that the logically entailed conclusion and an implicature have equal footing. In fact, Grice conceived of implicatures as additional pieces of information which may vary in their salience for the hearer. Second, most syllogisms,especially those that have existential conclusions, are quite difficult and the processing required to draw implicatures (as well as valid conclusions) might well be inaccessible

in such tasks. Third, a critical feature of Grice's theory is that implicatures are context-dependent and not always carried by a given utterance in all contexts. Gricean implicatures are considered defeasible and sensitive to context. Syllogistic reasoning, in particular, encourages participants to constrain their conclusions and to respond minimally to the conclusion posed (or prompted).

[4] Two other relevant Gricean properties are calculability and non-conventionality. Calculability is the idea that one should be able to reconstruct the argument that led to an implicature. Non-conventionality is the idea that implicatures are not part of the conventional meaning of a term (See Levinson, 1983).

[5]To appreciate a particularized implicature, consider the following exchange.

Al: Is the President around?

Hillary: Well, his jacket is in the hallway.

Hillary's response does not answer the yes/no question directly. However, one can work out the Gricean implicature (yes). Note that there is no general rule (like one for weak scalar terms) that is linked to a specific word.

References

Acredolo, C. & Horobin, K. (1987). Development of relational reasoning and premature closure. Developmental Psychology, 23, No. 1, 13-21.

Braine, M. D. S., O'Brien, D. P. Noveck, I. A., Samuels, M., Fisch, S. M., Lea, R. B., and Yang, Y. (1995). Predicting intermediate and multiple conclusions in propositional logic inference problems: Further evidence for a mental logic. Journal of Experimental Psychology: General, 124, No. 3, 263-292.

Braine, M. & Rumain, B. (1981). Children's comprehension of "or": Evidence for a sequence of competencies. Journal of Experimental Child Psychology, 31, 46-70.

Braine, M. D. S., & Rumain, B. (1983). Logical Reasoning. In P. Mussen (Ed.), Handbook of Child Psychology, Vol. 3. New York: Wiley.

Byrnes, J. P. & Overton, W. F. (1986). Reasoning about certainty and uncertainty in concrete, causal, and propositional contexts. Developmental Psychology, 22, 793-799.

Carston, R. (1998). Informativeness, relevance, and scalar implicature. In Carston, R. and Uchida, S. (eds.). Relevance theory: Applications and implications. John Benjamins.

Chierchia, G. (1999). Scalar implicature as a Polarity phenomenon. Proceedings from International Symposium on Linguistics (Economy in Language Design, Computation and Use). Lyon, France.

Gazdar, G. A. (1979). Pragmatics: Implicatures, Presupposition and Logical Form. New York: Academic Press.

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

Hilton, D. (1995). The social context of reasoning: Conversational inference and rational judgment. Psychological Bulletin, 118 (No. 2), 248-271.

Horn, L. (1973). Greek Grice. Proceedings of Ninth Regional Meeting of the Chicago Linguistic Society, 205-214.

Johnson-Laird, P. N., & Bara, B. (1984). Syllogistic Inference. Cognition, 16, 1-61.

Levinson, S. (1983). Pragmatics. Cambridge: Cambridge University Press.

Newstead, S. (1995). Gricean implicatures and syllogistic reasoning. Journal of Memory and Language, 34. 644-664.

Noveck, I. A., Ho, S. & Sera, M. (1996). Children's understanding of epistemic modals. Journal of Child Language, 23, 3, 621-643.

Paris, S. (1973). Comprehension of language connectives and propositional logical relationships. Journal of Experimental Child Psychology, 16, 278-291.

Politzer, G. (1986). Laws of language use and formal logic. Journal of Psycholinguistic Research, Vol. 15, 47-92.

Politzer, G. (1991). Immediate deductions between quantified sentences. In Lines of Thinking: Vol. 1. Gilhooley K. J ., Keane, M. T. G. & Logie, R. H. & Erdos, G. (eds.). London: John Wiley.

Politzer, G. & Noveck, I. A. (1991). Are conjunction rule violations the result of conversational rule violations? Journal of Psycholinguistic Research, Vol. 20, No. 2, 83-103.

Shipley, E. F. (1979). The class-inclusion task: Question form and distributive comparisons. Journal of Psycholinguistic Research, 8, No. 4, 301-331.

Smith, C. L. (1980). Quantifiers and question answering in young children. Journal of Experimental Child Psychology, 30, 191-205.

Sophian, C. & Somerville, S. (1988). Early developments in logical reasoning: Considering alternative possibilities. Cognitive Development, 3, 183-222.

Sperber, D. & Wilson, D. (1985/1996). Relevance: Communication and Cognition. Oxford: Basil Blackwell.

Stedmon, J. A. (1986). More than 'All'? Children's problems with plural judgements. In Reasoning and Discourse Processes. I. Myers, K. Brown, B. McGonigle (eds.). London: Academic Press.

Table 1

A truth table representation of the task in Experiments 1 and 2.

Statement about covered box	Parrot only box	Parrot + Bear box	Evaluation of statement
There has to be a parrot	True	True	True
There does not have to be a parrot	False	False	False
There might be a parrot*	True	True	True
There cannot be a parrot	False	False	False
There has to be a bear	False	True	False
There does not have to be a bear	True	False	True
There might be a bear	False	True	True
There cannot be a bear	True	False	False

Notes. Participants see two boxes, one that contains a parrot by itself and another containing a parrot and a bear. They are shown a third box that remains covered and told "All I (the experimenter) know is that this (Covered) box looks like this (Parrot-only) box or this (Parrot + Bear) box." All statements end with the phrase "in the box." * This statement underdetermines the actual state of affairs because there has to be a parrot in the covered box.

Table 2. Percentage of correct responses to modal statements which concerned what was necessarily in the hidden box (i.e., the parrot) and what was possibly in the hidden box (i.e., the bear) in Experiment 1.

		Age			
		5 (n=32)	7 (n=20)	9 (n=16)	Adults (n=20)
Presented statements	Is the puppet right?	Necessary Conclusion (parrot)			
Has to be a parrot	yes	75*	90**	88**	100**
Does not have to be a parrot	no	72*	75*	75*	100**
Might be a parrot	yes	72*	80**	69	35
Cannot be a parrot	no	66	80**	100**	100**
Total		73**	81**	83**	83**
		Possible Conclusion (bear)			
Has to be a bear	no	47	65	88**	100**
Does not have to be a bear	yes	66	75*	81**	100**
Might be a bear	yes	53	80**	100**	100**
Cannot be a bear	no	53	80**	100**	100**
Total		55	75**	92**	100**

Notes. * $p < .05$, ** $p < .01$.

Table 3. Participants' reponses to two sets of modal statements after intensive training in Experiment 2.

Evaluations of Necessary Conclusion

		Age		
		5	7	Adults
Presented Statements	Is the puppet right?	(n=16)	(n=16)	(n=16)
Must be a parrot	yes	81*	100**	94**
Might not be a parrot	no	75*	94*	94**
Might be a parrot	yes	81*	94**	75*
Must not be a parrot	no	69	94**	100**
Total		77	95	91
Has to be a parrot	yes	94**	100**	100**
Does not have to be a parrot	no	63	94**	100**
Could be a parrot	yes	88*	100**	94**
Could not be a parrot	no	75*	100**	100**
Total		80	98	98

Table 3 continued.

Evaluations of Possible Conclusion

		Age		
		5 (n=16)	7 (n=16)	Adults (n=16)
Must be a bear	no	31	88*	100**
Might not be a bear	yes	50	94**	100**
Might be a bear	yes	88*	88*	100**
Must not be a bear	no	31	56	94**
Total		52	81	98
Has to be a bear	no	37	69	100**
Does not have to be a bear	yes	63	81*	100**
Could be a bear	yes	75*	94**	100**
Could not be a bear	no	37	50	50
Total		53	73	88

Notes. * $p < .05$, ** $p < .01$. These modal statements were presented in two separate sets. The statements in the top half of the necessity set were presented in a random order along with the statements in the top half of the possibility set; likewise for the statements in the two bottom halves.

Table 4. Rates of correct responses to the six types of sentences presented in Experiment 3.

Sentence Type	Correct Response	Age		
		7-8	10-11	Adults
Utterances expressed with all		(n=31)	(n=30)	(n=15)
Absurd (false)	No	93%	99%	99%
(e.g. All chairs tell time)				
Appropriate (true)	Yes	91%	99%	96%
(e.g. All elephants have trunks)				
Inappropriate (false)	No	86%	99%	96%
(e.g. All dogs have spots)				
Utterances expressed with some				
Absurd (false)	No	95%	99%	98%
(e.g. Some stores are made of bubbles)				
Appropriate (true)	Yes	84%	90%	99%
(e.g. Some birds live in cages)				
Inappropriate (true though pragmatically infelicitous)	Yes	89%	85%	41%
(e.g. Some giraffes have long necks)				

Notes. There were five exemplars of each type.

Appendix

The materials in Experiment 3 (translated from French).

Bizarre	Factually universal	Factually existential
All birds have telephones.	All books have pages.	All dogs have spots.
All crayons have noses.	All hammers have a handle.	All books have color pictures.
All chairs tell time.	All robins have wings.	All animals are striped.
All garages sing.	All elephants have trunks.	All clothes have zippers.
All couches have windows.	All refrigerators have doors.	All birds live in cages.
Some fish are made of leaves.	Some televisions have screens.	Some flowers are yellow.
Some fruits have computers.	Some cars have motors.	Some dresses have pockets.
Some books are good to eat.	Some giraffes have long necks.	Some tools are made of wood.
Some stores are made of bubbles.	Some cats have ears.	Some children are blond.
Some children are made of feathers.	Some airplanes have wings.	Some drinks are made from chocolate.

Notes. Underlined statements in the middle column are underdeterminate. In order to prepare multiple sets of materials, the contents of the top half of the last two columns were systematically switched with the those in their bottom half (e.g. All elephants have trunks becomes Some elephants have trunks and Some dresses have pockets becomes All dresses have pockets). Also, order of (randomized) All and (randomized) Some sentences was varied.