

Journal of Semantics, 2007, 24(4), 331-343
Oxford Journals doi:10.1093/jos/ffm006

The Psychological Reality of Classical Quantifier Entailment Properties

Guy Politzer

Institut Jean Nicod, C.N.R.S., Paris

Abstract

A test of directional entailment properties of classical quantifiers defined by the theory of generalised quantifiers (Barwise and Cooper, 1981) is described. Participants had to solve a task which consisted of four kinds of inference. In the first one, the premise was of the type *Q - hyponym - verb - blank predicate*, where Q is a classical quantifier, (e. g., *some cats are []*), and the question was to indicate what, if anything, can be concluded by filling up the slots in *.....- hyperonym - verb - blank predicate* (e. g., *.....animals are []*). The second kind of inference was the same, except that the hyperonym was in the premise and the hyponym in the conclusion. The third and fourth kinds of inference differed from the first two by the position of the hyperonym (resp. hyponym) which occupied the place of the predicate (e. g., *some [] are animals*). It was observed that when the directional entailment holds people respond accordingly in most cases and that when the entailment does not hold they correctly fail to produce it. These results provide elementary, but essential empirical support to this semantic approach to quantification, as well as a prerequisite for its application to the study of reasoning with quantifiers. The implications for the psychology of reasoning are discussed.

The Psychological Reality of Classical Quantifier Entailment Properties

1. Introduction

Generalised quantification theory (Barwise & Cooper, 1981) posits that a quantifier is a set of subsets of a domain. This approach has offered a number of important concepts to semantic theory. One of them is *directional entailment* (also called *monotonicity*). Informally, this concept captures the intuition that, for example, *all dogs* entails *all spaniels*, or that *some spaniels* entails *some dogs*. Formally, a quantifier Q_D is *downward entailing* if, for all domains D : $Q_D Y$ and $X \subseteq Y \subseteq D$ entails $Q_D X$. This is instantiated by the former example. Similarly, a quantifier Q_D is *upward entailing* if, for all domains D : $Q_D X$ and $X \subseteq Y \subseteq D$ entails $Q_D Y$ (this is instantiated by the latter example). Directional entailment can also be defined for binary relations between subsets (see Westerståhl, 2001). $Q_D XY$ (with X subject and Y predicate) is *left downward entailing* if, for all D : $Q_D XY$ and $X' \subseteq X \subseteq D$ entails $Q_D X'Y$ (for instance *all dogs bark* entails *all spaniels bark*); it is *right downward entailing* if, for all D : $Q_D XY$ and $Y' \subseteq Y \subseteq D$ entails $Q_D XY'$ (*no amphibians are dogs* entails *no amphibians are spaniels*). Analogous definitions obtain for left and right upward entailment, *mutatis mutandis*.

Viewed in these terms, the four classical quantified sentences have the following entailing characteristics (using the standard notations A, I, E, and O):

A (all X are Y): left downward, e. g., on an appropriate domain, *all antiques are vases* entails *all antiques for sale are vases*, and right upward, e. g., *all antiques are Ming vases* entails *all antiques are vases*.

I (some X are Y): left upward, e. g., *some antiques for sale are Ming vases* entails *some antiques are vases*, and right upward (like *all*).

E (no X are y): left downward (like *all*) and right downward, e. g., *no antiques are vases* entails *no antiques are Ming vases*.

O (some X are not Y): left upward (like *some*) and right downward (like *no*).

Over the last two decades the concept of directional entailment has played an important role in semantic theory, e.g. in accounts of the distributional properties of negative polarity items (Ladusaw, 1980) and the discussion of scalar implicature computation (these implicatures are suspended in downward environments: Horn, 1989; Chierchia, 2004).

From the viewpoint of generalised quantifiers in semantic theory, an individual who knows the meaning of a quantifier must have the ability to carry out the directional

entailments. If people are indeed proficient at this, then this skill might go some way towards explaining people's reasoning with quantified sentences. Hence, Geurts (2003) proposed that these entailment properties are highly relevant to reasoning, and in particular to syllogistic reasoning. But psychologists unfamiliar with this conceptual framework may be puzzled, if not sceptical, with regard to the psychological reality of entailment relations that stem from work in philosophical logic and have been endorsed only on the basis of linguistic intuition. Moreover, the notion that reasoning might be taken care of by formal properties of the semantic component of language rather than relying on some logical system (rule-based or model-based) is totally alien to most psychologists working in the field. A first step towards getting the entailment properties on the psychologists' agenda would be to ascertain that people do master these. So far, linguists have not felt it necessary to experimentally investigate the extent to which people do so, presumably because the existence of entailment properties of quantifiers is not in dispute. Geurts (2003, Geurts & van der Slick, 2005) did carry out experiments on quantifier entailment but the aim of these investigations was more elaborate: In the former case he tested an hypothesis about the relative complexity of various quantifiers, and in the latter he used fairly complex sentences containing two quantifiers and limited to right entailment. In brief, there is still a need for an experimental demonstration that people master the basic entailment arguments described earlier, which is an elementary assumption to make if generalised quantification theory provides an adequate framework for the representation of, and reasoning with, quantifiers. The work that will be reported is but a first step limited to the quantifiers that have been studied most by psychologists, namely the classical Aristotelian quantifiers; there is, however, no reason why the method used could not be applied to other quantifiers, especially to those whose entailment properties are intuitively less clear.

2. Experiment

2.1. Materials and design

The experiment was administered in written form. Four types of one-premise arguments were prepared. They will be described by way of an example, where the relational term *are* is used, and in which Q stands for one of the four classical quantifiers.

Type 1. Premise: Q animals are []
 Conclusion: cats.....[]

This type is defined by the following two characteristics: (i) the "subject" (i. e., the restriction) is provided, but the predicate (here in square brackets) which is common to both sentences, is left unspecified; and (ii) the conclusion's subject is a hyponym of the premise's subject. The two dotted slots were prepared to accommodate the participant's response, namely a quantifier and a verbal expression (is/ is not/ are/ are not), respectively. An example of expected correct inference, with the quantifier *all*, would be:

Premise: all animals are []
 Conclusion: all cats are []

This is a test of left downward entailment.

Type 2. Premise: Q [] are animals
 Conclusion:[]cats

In this type, (i) the common subject in square brackets is left unspecified, and (ii) the conclusion's predicate is a hyponym of the premise's predicate. The dotted slots serve the same purpose as previously. This type tests for right downward entailment.

Type 3. Premise: Q cats are []
 Conclusion:animals.....[]

This is a variant of type 1: It differs only in the hyponymy relation which now is reversed and serves to test left upward entailment.

Type 4. Premise: Q [] are cats
 Conclusion:[]animals

Similarly, type 4 differs from type 2 by the reversal of the hyponymy relation and serves to test for right upward entailment.

The four types appeared with all four quantifiers each, hence 16 different arguments which will be designated by the traditional letter (A, I, E, or O) and a number referring to the type. The example given above for type 1 is an A1 argument. Out of the 16 arguments, eight are entailing and the other eight are non-entailing. The present study focused on people's recognition of quantifiers' entailment properties, so it is the first eight arguments that specifically address this question. Their eight non-entailing counterparts (A3 for A1, E3 for E1, etc.) were also considered because they provide complementary evidence in the form of a control: If people produce a conclusion for, e.g., an A1 argument showing that they recognise a left downward entailment from *all animals are []* to *all cats are []*, they should *not* produce a conclusion for the associated A3 argument from *all cats are []* to *all animals are []*

because this would entail that they execute a left upward entailment at the same time as they do the left downward entailment.

Unspecified subjects and predicates were used to make the task as context neutral as possible. This was to avoid that the premises be associated, or restricted, to specific categories or attributes that could affect participants' inferences for emotional or cognitive reasons, and so to enable one to generalise the results safely. To make the task natural, the predicates in argument-types 1 and 3 were not presented as empty spaces in brackets as above, but were occupied by a pronoun. As the experiment was administered in French, it was easy to take advantage of the very natural way that this language refers to an attribute already present in the context by use of a pronoun (e.g., *all cats are []* is rendered by *quelques chats le sont*, where the neutral pronoun *le* occupies a place that is empty in English). So, participants were told that the pronoun referred to some unspecified attribute which they need not be concerned about. Similarly, and even more straightforwardly, for the eight argument-types 2 and 4, it was said in the instructions that the premises referred to some set of objects (the nature of which, again, participants need not be concerned about). A premise such as *all of them are cats* was naturally formulated as *all are cats* (French *tous sont des chats*) without a need for a pronoun (*them*).

As performance could be influenced by the relational term that was used (*are*), the arguments were also formulated with *have*, so that the previous examples in this second version became: *all cats have []*, *all [] have cats*, etc. Again there was no specific predicate to occupy the empty slot as it is very natural in French to use the pronoun *en* (e.g., *tous les chats en ont*, *all cats have*).

Finally, French has a very common way to express the particular quantifier, namely *il y a*. (In English, its closest equivalent *there is/there are* occurs less frequently and cannot be used to paraphrase *some*-sentences as routinely as *il y a* for *quelques*-sentences in French). This provided an additional condition to generalise the results beyond the case of one single way to express particular quantification.

Four semantic domains were used: animals, vegetables, flowers, and fruit, with four different and very common hyponyms in each case (e.g., cat, dog, horse, goat), hence 16 pairs of hyponyms-hyperonyms; each pair was randomly allocated to one argument, with the constraint that the four semantic domains appear in each argument type. A sample of four items is given in the Appendix.

The arguments were presented in booklets prefaced with the instructions. The eight arguments of type 1 and 3 constituted one block, as did the eight arguments of

type 2 and 4. Within a block, the arguments were rotated with the constraint that no two identical quantifiers could be adjacent, and that the types should alternate. One block was presented first to half of the participants, and the other was presented first to the other half. In brief, the design was a 2 (verbal form of relation: *are vs have*) x 2 (formulation of the particular quantifier) x 2 (blocks) x 16 (argument type) factorial design, with repeated measures on the last two factors.

2.2. Participants and procedure.

The participants were 89 second year psychology students in the University of Paris-8. All were native speakers of French and untutored in logic. The experiment was administered during classes.

Participants were instructed to assume that the premises were true; they were then asked to enter into blank spaces (i) a quantifier that would be guaranteed to make the conclusion surely true; (ii) the appropriate verbal expression (*are /are not /is /is not* or the analog with *have*, depending on the experimental condition). They were told that it was not always possible to find a conclusion that was surely true; whenever they thought this applied, they had to place a cross in the empty spaces of the conclusion. In all cases, they were asked to indicate how confident they were about their answers on a four-point scale ranging from zero (not sure at all) to 3 (very sure). A high level of certainty should accompany the identification or the application of a formal argument, whereas a lower level of certainty should be observed when the response follows from the absence of such an identification. Previous research on deduction with quantified sentences has established the reliability and validity of this measurement (Politzer, 1991).

2.3. Results

There was no effect of order and no effect of the formulation of the quantifier. There was a significant effect of the *are vs. have* factor for only one of the 16 argument forms, to be discussed later. In brief, none of the three binary factors had an effect that was both systematic and significant, so the results were pooled across these factors. For any of the arguments, a judgment of entailment was attributed to participants when the quantifier they entered in the conclusion reproduced that in the premise (and similarly for the relational term). Non-entailment was operationally defined by any answer showing the absence of the reproduction of the premise quantifier, in line with the logical definition of entailingness. Take for example *all flowers are [] thereforeroses are []*. Participants were credited with a left downward entailing answer if they filled in the dotted space with *all*. They were considered as giving a non-entailing answer in the

other cases, that is, if (i) they indicated that nothing followed; or (ii) gave a different quantifier, whether logically incorrect (e. g., *no*) or correct (e. g., *some*). Of course in the latter case, giving a correct conclusion is an indication that the participant masters at least some of the semantic content of the quantifier in question; in addition, even though such an answer is preferred to the entailing answer, it is not incompatible with its recognition. However, because a conservative approach was adopted, these answers were discounted. In fact, they turned out to be very rare (13 occurrences for the four arguments where they could occur, that is less than 4 percent of the answers).

Table 1 displays for each of the 16 argument forms the percentages of entailing answers and (by complementarity) of non-entailing answers. The pattern of results is very clear.

Table 1. Percentage of responses and mean certainty ratings associated with the correct response for the 16 argument forms.

Response	↓A1	↓E1	E2↓	O2↓	↑I3	↑O3	A4↑	I4↑
entailing°	88.8	86.4	83.1	43.7	92.2	81.8	93.3	71.9
non-entailing	11.2	13.6	16.9	56.3	7.9	18.2	6.7	28.1
mean certainty	2.68	2.63	2.66	2.32	2.40	2.28	2.65	2.50

Response	A3	E3	E4	O4	I1	O1	A2	I2
entailing	8.0	11.2	16.9	35.6	33.7	28.1	9.0	47.2
non-entailing°	92.0	88.8	83.1	64.4	66.3	71.9	91.0	52.8
mean certainty	2.32	2.21	1.99	1.87	1.88	1.94	2.14	2.09

Entailing quantifiers are marked with an arrow indicating side and direction

° correct responses

First, for the eight entailing cases, all except O2 were identified as such. The comparison of the percentages for and against entailment in each case shows the magnitude of the preferences: The weakest significant majority was 72% (for I4) and the strongest 93% (for A4). These differences were all significant at the .01 level (chi-square test). The O2 case did not elicit any significant majority and constitutes an exception. Pooling the eight arguments together, the mean frequency of entailing responses was 80% (86% excluding the O2 case). A similar pattern obtained for the

eight non-entailing arguments¹. All except I2 elicited a strong majority against entailment (weakest: 64% for O4, strongest: 92% for A3) and the associated differences were all significant (at the .05 level, most being well beyond this level). The mean frequency of non-entailment responses for these arguments was equal to 76% (80% excluding I2). In brief, on 14 out of 16 arguments, participants were strongly sensitive to the presence or to the absence of the entailment property of the four classical quantifiers, while two arguments, both of type 2, and both particular, seem to have remained opaque to participants.

3. Discussion

With two exceptions (O2 and I2, which will be considered later), participants' pattern of inferences conformed to the expected pattern determined by directional entailments, in strong support of the hypothesis. Interestingly, the analysis of the certainty ratings shows a pattern in perfect agreement with a well-known phenomenon in the psychology of reasoning, namely that individuals are better at identifying valid than invalid arguments. For the entailment answers to entailing arguments, the certainty ratings ranged from 2.28 (for O3) to 2.68 (for A1) with a mean equal to 2.52, whereas for the non-entailment answers to non-entailing arguments they ranged from 1.87 (for O4) to 2.32 (for A3) with a mean equal to 2.06. In brief, all but one of the entailing inferences were preformed with greater certainty than the non-entailing ones; in addition, the difference between the means is considerable, as it is almost half a unit on a scale whose range is equal to three units. To the extent that a high certainty of the response correlates with the operation of a rule or to the identification of a property, this result adds support to the notion that in solving the arguments participants demonstrated that they their recognised the entailment properties of the quantifiers.

The first exception to the overall pattern is the I2 argument; this should not be viewed as problematic for two reasons. One, although it indicates that participants who recognised its non-entailing property failed to constitute a significant majority, its counterpart I4 indicates that the majority correctly recognised the upward entailment. Now, it is not too surprising that the relatively low performance on the *valid* argument I4 (actually the poorest apart from O3) is accompanied with an even poorer performance on the associated *invalid* I2 argument for reasons that have just been mentioned. Two, the present result for the I2-I4 pair might be not entirely robust as it

¹ The expression "non-entailing" is used to qualify one member of a pair of arguments such as A1/A3; this would be improper if used to qualify the associated quantifier, which is always entailing.

was confirmed in only one of two unpublished replication studies run by the author; interestingly, the failure to replicate concerns a sample of high school students with a firmer background in mathematics, suggesting (tentatively) a relatively greater difficulty of this pair, although insufficient to affect participants with higher logical abilities. In terms of difficulty, it is noteworthy that performance was, by far, worse for particular than for universal quantifiers, both in terms of certainty ratings and of entailing and non-entailing correct answers: Universals were the four most often correctly identified as non-entailing, and they ranked 1st, 3rd, 4th and 5th among the most often correctly identified as entailing. This result may be understood in set-theoretical terms as reflecting that universal quantifiers *all* and *no* are more simple than their particular counterparts²: the greater formal complexity results in a greater processing load.

The other exception, the O2 argument, is also a particular argument. It contradicts the hypothesis more directly as it is entailing; in other words, the right downward entailment was not recognised for the O sentence. If we notice that the associated O4 argument showed only a weak majority against upward entailment, it would seem that a number of individuals were confused with regard to the direction (upward or downward) of right entailment for this specific quantifier. This conclusion is reinforced by the results of the two replication studies, both of which yielded percentages similar to the present experiment. The result for O2 supports Geurts's claim that negation adds cognitive load to the processing of particular quantifiers (which has been documented for a long time, see, e. g., Anderson, 1981). That no similar effect was observed for the other tests of the *some...not* sentence, namely the O1 and O3 arguments, which concern left (subject) entailment, is consistent with the negation hypothesis, as negation in two-place relational quantifiers is linked to the predicate, not to the subject³. Also, the fact that the certainty ratings for O2 and O4 are among the lowest of the entailing and the non-entailing arguments, respectively, attests to the difficulty of this quantifier, while showing the coherence of the results. However, it could be objected that if the origin of the difficulty of O2 were due to negation, a similar difficulty for the other right entailing negative quantifier, namely *no* tested by E2 should have been observed. But this objection is not compelling because considering the greater ease of processing of universals mentioned above, an interaction between

² This can be illustrated with "number triangles" which show that *all* and *no* occupy just one edge of the triangle whereas their counterparts occupy the whole complement of the domain.

³ Left entailment is a property attached to the subject, which makes a partition between universal and particular sentences (odd arguments), whereas right entailment is a property attached to the predicate, which makes a partition between affirmative and negative sentences (even arguments).

quantity and quality can be expected; that is, it is reasonable to assume that the additional processing load due to negation has a compounding effect for particulars but not for universals.

It was mentioned earlier that, out of the 16 arguments, there was only one significant difference between the *are* and the *have* formulations of the sentences. Interestingly, this concerns the O2 argument, for which the rate of correct responses was about twice as high with *have* as it was with *are*. That failure on the O2 argument occurred with one verb but not with the other, while the same manipulation did not affect the other arguments, speaks in favour of the idea that this argument has some peculiarities. One could think of the effect of a scalar implicature. What such an effect could consist of has to be worked out on a case by case basis for each particular premise. It is easy to see that, should a scalar implicature be triggered, this could have no effect for six of the arguments, including O2. For O2, given, for example, *some [] are not animals*, one can expect that participants who rightly infer *some [] are not cats* will carry out the same inference in case they add the scalar implicature *some [] are animals* to the premise, because this implicated premise does not interfere with the entailment property of the explicit premise. In brief, a scalar implicature cannot justify the failure to produce the right response for O2. It can be verified that this applies also, *mutatis mutandis*, to O3, I3 and I4 (which are entailing) and to O1 and I1 (which are not). There are, however, two non-entailing cases, I2 and O4, that, in principle, could be affected by an implicated premise. Consider I2: Participants who correctly fail to draw the conclusion *some [] are cats* from *some [] are animals* are justified to conclude *some [] are not cats* if they have added the implicated premise *some [] are not animals*, because this turns I2 into the downward entailing argument O2. The same applies to the non-entailing O4 whose implicated premise turns it into the upward entailing argument I4. Now, a scrutiny of the raw data indicates that such responses are very rare (6% of the responses to I2 and O4), from which it can be concluded that implicatures do not affect the performance in the present task; this rules out, in particular, a possible explanation for the deviant performance on O2.

One cannot underestimate the theoretical importance of the entailments investigated here, even if they might appear easy enough to carry out. In fact, that these inferences seem to be so easy is part of the explanandum: the fact that specific inferences are executed so readily highlights that they are critical for people's understanding of quantifiers as predicted by the theory. Also, the entailment task was designed as a precise test and, in view of the number of argument forms tested, a

severe test of generalised quantification theory. The test can also be considered as severe from the point of view of the task demands, as the format of response required that participants produce their own conclusion rather than evaluate a correct conclusion offered to them, which might have been too transparent. That the theory has stood up well (and independently of any expectations one might have about the results of the test based on intuition) lends strong support to it. Of course, the conclusion that people recognise most entailment properties of the quantifiers considered here obviously holds to the extent that participants have been submitted to a valid test, that is, to a task containing arguments that faithfully operationalise the formal definition of quantifier entailment. It does not seem questionable that this has been done in the present experiment, as (i) the premises presented the quantified sentences of interest; (ii) the inclusion relations could not fail to be understood as such, as they were chosen from highly familiar categorisations; and (iii) the conclusions disclosed directly whether or not the expected entailed quantified sentences were produced. This suggests that the task could be exploited to investigate non-classical quantifiers as well.

Finally, the present argument forms may be described as syllogisms in disguise: Indeed, one can associate a syllogism containing an implicit A minor premise with each argument form (e. g., EA1 to E1, etc.) Consequently, the claim that the performance on the task demonstrates people's mastery of the entailment properties of classical quantifiers justifies another claim, namely that people can solve this kind of syllogism by whatever means they possess to solve syllogistic tasks in general. This objection must be considered seriously; but to substantiate it, supporters of the relevant theories would have to apply them to explain the present data, including data that have not been reported in the results section.⁴ This concerns a few striking differences in performance on several arguments when compared with performance on standard syllogisms. To take one example, the A3 argument yielded 8% of A conclusions, whereas the usual rate of A conclusions for the AA3 syllogism is about 65%. More important, it seems that all theories of syllogism, whether they are psychological or were developed in classical logic, incorporate some principles that are indistinguishable from one or another entailment property. This claim cannot be developed in the present paper, but two examples will suffice: The *dictum de omni et nullo*, which is at the basis of the medieval theories of the syllogism, can be shown to be a mere consequence of

⁴ Of course, a theory based on the entailing properties of quantifiers faces the same challenge: some logical principle would have to be added to such a theory to account for the detailed responses offered by participants.

entailment properties (Hoeksema, 1986). Similarly, the concept of distribution used in the formulation of the classic laws of the syllogisms can be reformulated in terms of downward entailingness (Makinson, 1969). The upshot of all this is that it is not unreasonable, based on considerations of parcimony and generality, to view the quantifier entailment properties as primitive and reasoning principles as derived.

The present investigation concerns itself with classically quantified sentences. Quantifier entailing properties are highly relevant to the psychology of reasoning. As Gueurts's (2003) study exemplifies, monotonicity is a plausible main component of a theory of reasoning with quantified sentences – more specifically of syllogistic reasoning. Now, the relevance of monotonicity for reasoning can be extended far beyond quantification, as the various syntactic categories of sentences have monotonicity properties (van Benthem, 1986; Sanchez-Valencia, 1995). That is, given two expressions of the same type such that $[A] \subseteq [B]$, inference in natural language can be viewed as a matter of substituting A with B (or B with A) once the monotonicity of the syntactic category concerned is determined (which requires a procedure of polarity marking). There follows the idea of a general theory of deductive reasoning in natural language based on monotonicity properties of grammatical components of sentences. Because monotonicity properties are alien to the current theories of deductive reasoning, whether these are based on mental models (Johnson-Laird & Byrne, 1999), or on mental rules (Rips, 1994; Braine & O'Brien, 1998) or on a more recently developed probabilistic point of view (Chater & Oaksford, 1999; Oaksford, 2007), for psychologists supporting any one of them, the grammatical approach may appear as a revolutionary perspective, certainly worth considering closely.

Acknowledgment

The author thanks Bart Geurts for his comments on a first draft of this paper

Appendix

Four test items with an entailing quantifier in each of the four types are given below in the original French version (with their English translation in the right column).

↓A1

tous les animaux le sont.

donc:

_____ chien(s).....

all animals are.

therefore:

_____ dogs(s).....

↑O3

quelques tulipes ne le sont pas.

donc:

_____ fleur(s).....

some tulips are not.

therefore:

_____ flower(s).....

E2↓

aucun n'est un légume.

donc:

_____.....poireau(x)

none are vegetables.

therefore:

_____.....leek(s)

I4↑

quelques-uns sont des oranges.

donc:

_____..... fruits

some are oranges.

therefore:

_____..... fruit

References

- Anderson, John R. (1981), Memory for logical quantifiers. *Journal of verbal Learning and Verbal Behavior* **20**: 306-321.
- Barwise, Jon, & Robin Cooper (1981), Generalized quantifiers and natural language. *Linguistics and Philosophy* **4**: 159-219.
- Braine, Martin. D. S. , & David P. O'Brien (1998), *Mental logic*. Lawrence Erlbaum. Mahwah, N. J.
- Chater, Nick. , & Mike Oaksford (1999), The probability heuristics model of syllogistic reasoning. *Cognitive Psychology* **38**: 191-258.
- Chierchia, Gennaro (2004), Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (ed.), *The cartography of syntactic structures*. Vol 3: Structures and beyond. Oxford University Press. Oxford.
- Geurts, Bart (2003), Reasoning with quantifiers. *Cognition* **86**: 223-251.
- Geurts, Bart, & Frans van der Slick (2005), Monotonicity and processing load. *Journal of Semantics* **22**: 97-117.
- Hoeksema, Jack (1986), Monotonicity phenomena in natural language. *Linguistic Analysis* **16**: 25-40.
- Horn, Laurence R. (1989), *A natural history of negation*. University of Chicago Press. Chicago.
- Johnson-Laird, Philip N. , & Ruth M. J. Byrne (1991), *Deduction*. Lawrence Erlbaum. Hove & London.
- Ladusaw, William (1980), *Polarity Sensitivity as Inherent Scope Relations*. Garland Press. New York.
- Makinson, David (1969), Remarks on the concept of distribution in traditional logic. *Noûs* **3**: 103-108.
- Oaksford, Mike , & Nick Chater (2007), *Bayesian rationality: the probabilistic approach to human reasoning*. Oxford University Press. Oxford.
- Politzer, Guy (1991), Immediate deduction between quantifiers. In Kenneth J. Gilhooly, Mark T. G. Keane, Robert H. Logie, & George Erdos (eds.), *Lines of Thinking*. Vol. 1. Wiley. London. 85-97.
- Rips, Lance J. (1994). *The psychology of proof*. M. I. T. Press. Cambridge, Ma.
- Sánchez-Valencia, Victor (1995), Parsing-driven inference: Natural logic. *Linguistic Analysis* **25**: 258-285.
- van Benthem, Johan (1986), *Essays in logical semantics*. Reidel. Dordrecht.

Westerståhl, Dag (2001), Quantifiers. In Lou Goble (ed.), *Philosophical Logic*.
Blackwell. Oxford. 437-460.