

Truth value judgments vs. validity judgments

Elizabeth Coppock
SCAS, Uppsala University & University of Gothenburg

Texas Linguistics Society 2014

Outline

- 1 Introduction
- 2 Theories
- 3 Truth value judgments
- 4 Validity judgments

Guess the original version

A study of more than 2,500 retired NFL players found that those who had $\left\{ \begin{array}{l} \text{more than two} \\ \text{at least three} \end{array} \right\}$ concussions during their careers had triple the risk of clinical depression as those who had no concussions.

Guess the original version

A study of more than 2,500 retired NFL players found that those who had $\left\{ \begin{array}{l} \text{more than two} \\ \text{at least three} \end{array} \right\}$ concussions during their careers had triple the risk of clinical depression as those who had no concussions.

Guess the original version

The maximum number of cards in an Extra Deck is 15, although it is allowed to have $\left\{ \begin{array}{l} \text{fewer than 15} \\ \text{at most 14} \end{array} \right\}$ in it.

Guess the original version

The maximum number of cards in an Extra Deck is 15, although it is allowed to have $\left\{ \begin{array}{l} \text{fewer than 15} \\ \text{at most 14} \end{array} \right\}$ in it.

Guess the original version

I hardly ate $\left\{ \begin{array}{l} \text{more than 3} \\ \text{at least 4} \end{array} \right\}$ bites of the ham, cheese & egg lunch that
Les whipped up for us.

Guess the original version

I hardly ate $\left\{ \begin{array}{l} \text{more than 3} \\ \text{at least 4} \end{array} \right\}$ bites of the ham, cheese & egg lunch that
Les whipped up for us.

Guess the original version

Moreover, in these variables there are $\left\{ \begin{array}{l} \text{at least four} \\ \text{more than three} \end{array} \right\}$ fluctuation parameters, namely G^{pp} , G^{JJ} , G^{pJ} , $G^{p\bar{J}}$, G^{JJ} and $G^{\bar{J}\bar{J}}$, since we are using one complex variable J .

Guess the original version

Moreover, in these variables there are $\left\{ \begin{array}{l} \text{at least four} \\ \text{more than three} \end{array} \right\}$ fluctuation parameters, namely G^{pp} , $G^{J\bar{J}}$, G^{pJ} , $G^{p\bar{J}}$, G^{JJ} and $G^{\bar{J}\bar{J}}$, since we are using one complex variable J .

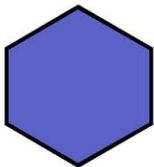
Guess the original version

I am in my sixties and had $\left\{ \begin{array}{l} \text{more than two} \\ \text{at least three} \end{array} \right\}$ parents growing up.

Guess the original version

I am in my sixties and had $\left\{ \begin{array}{l} \text{more than two} \\ \text{at least three} \end{array} \right\}$ parents growing up.

Superlative modifiers and ignorance



A hexagon has more than four sides.

#A hexagon has at least five sides.

(Nouwen 2010)

At least conveys ignorance; *more than* does not.

Validity judgment experiments

Liz had 3 beers \Rightarrow *Liz had more than 2 beers.* 100%

Liz had 3 beers \Rightarrow *Liz had at least 3 beers.* 50%

Liz had 3 beers \Rightarrow *Liz had fewer than 4 beers* 93%

Liz had 3 beers \Rightarrow *Liz had at most 3 beers* 61%

(Geurts et al. 2010)

Outline

- 1 Introduction
- 2 Theories**
- 3 Truth value judgments
- 4 Validity judgments

Two classes of theories about superlative modifiers

Ignorance as entailment (Geurts & Nouwen 2007)

- “Liz had at least 3 beers” is true if and only if *the speaker considers it necessary that Liz had 3 beers or more and considers it possible that Liz had more than 3 beers.*

Ignorance as implicature (Büring 2008, Cohen & Krifka 2011, Coppock & Brochhagen 2013b:i.a.)

- “Liz had at least 3 beers” is true if and only if *Liz had 3 beers or more.* The ignorance implication is an implicature.

Implicit disjunction theory

Büring (2008) (followed by Cummins & Katsos 2010 and Biezma 2013):

- *At least p* 'amounts to a disjunction' between *p* and *more than p*.
- There is an 'implicature schema' that says, if a speaker says *A or B*, then the speaker considers both *A* and *B* to be possible.

Implicit disjunction theory

Büring (2008) (followed by Cummins & Katsos 2010 and Biezma 2013):

- *At least p* 'amounts to a disjunction' between *p* and *more than p*.
- There is an 'implicature schema' that says, if a speaker says *A or B*, then the speaker considers both *A* and *B* to be possible.

Issue: In what sense is a speaker 'saying *A or B*' here?

Implicit disjunction theory

Büring (2008) (followed by Cummins & Katsos 2010 and Biezma 2013):

- *At least p* 'amounts to a disjunction' between *p* and *more than p*.
- There is an 'implicature schema' that says, if a speaker says *A or B*, then the speaker considers both *A* and *B* to be possible.

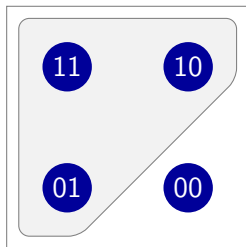
Issue: In what sense is a speaker 'saying *A or B*' here?

Twist on this view (Coppock & Brochhagen 2013b):

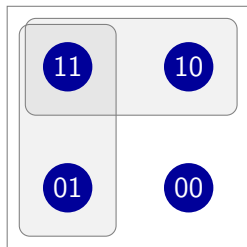
- Saying *at least p* is not saying '*p or more than p*'
- But *at least* and *at most* have an important property in common with disjunctions. Expressed using inquisitive semantics.

Traditional vs. Inquisitive Disjunction

Traditional disjunction

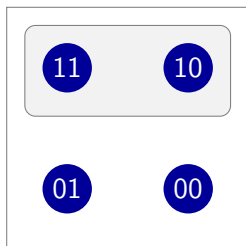


Inquisitive disjunction

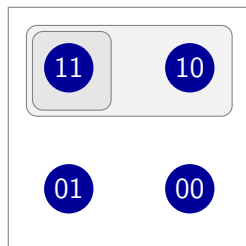


Inquisitive semantics analysis

Ann snores



At least Ann snores



at most vs. fewer than

Liz had *at most 4 beers* à la Coppock and Brochhagen:

- **logically implies** that Liz had fewer than 5 beers (like *Liz had fewer than 5 beers*)
- **brings up the issue** of whether Liz had fewer (unlike *Liz had fewer than 5 beers*)

“fewer than 5”



“at most 4”



Interactive sincerity (Coppock & Brochhagen 2013b)

Interactivity

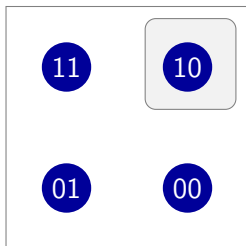
ϕ is **interactive** iff $\llbracket \phi \rrbracket$ contains more than one possibility.

Maxim of Interactive Sincerity

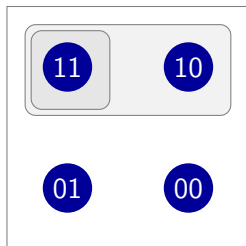
If ϕ is interactive, then ϕ is interactive in the speaker's information set.
'Don't bring up an issue that you already know how to resolve'

Example

Fred's information state



At least Ann snores



Fred should not assert *At least Ann snores*.

Related phenomena

Has the package arrived?

↪ The speaker considers it possible that it has arrived, and also considers it possible that it hasn't arrived.

My keys are either in my purse or in the car.

↪ The speaker considers it possible that her keys are in her purse, and also considers it possible that her keys are in the car.

Whatever she's cooking is delicious.

↪ The speaker doesn't know what she's cooking.

Outline

- 1 Introduction
- 2 Theories
- 3 Truth value judgments**
- 4 Validity judgments

Back to Geurts et al's results

Liz had 3 beers \Rightarrow *Liz had more than 2 beers.* 100%

Liz had 3 beers \Rightarrow *Liz had at least 3 beers.* 50%

Liz had 3 beers \Rightarrow *Liz had fewer than 4 beers* 93%

Liz had 3 beers \Rightarrow *Liz had at most 3 beers* 61%

(Geurts et al. 2010)

Validity as information delimitation

That damn Kaplan was promoted.

Therefore, Kaplan was promoted. (valid)

Kaplan was promoted.

Therefore, that damn Kaplan was promoted. (valid?)

Kaplan (1999): Logical validity is not about truth-preservation but rather about 'information delimitation': There must be no semantic information in the conclusion that is not already contained in the premises.

Surely these inferences are truth-preserving

Re: Geurts & Nouwen's (2007) proposal that superlative modifiers semantically encode speaker ignorance, Cohen & Krifka (2011) write:

Suppose John committed exactly four traffic violations, but nobody knows this, not even the police (who are the authority on the subject), and not even John himself. Then, it would still be truth that he committed at least three traffic violations, and these truth values depend only on what actually happened, not on anybody's beliefs.

Truth value judgment experiment



There are at least 6 dogs in the picture.

- True
- False

Truth value judgment experiment



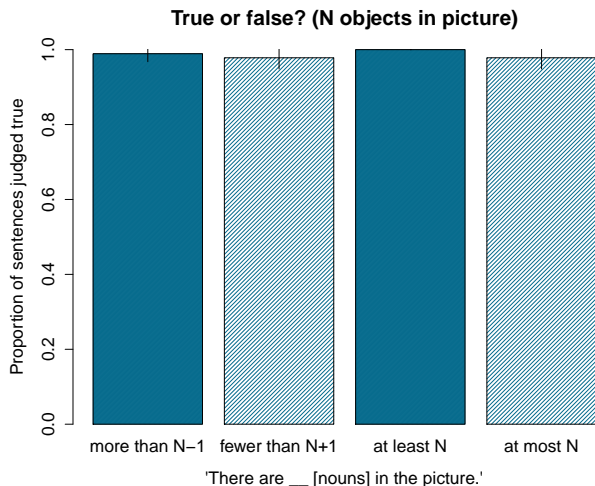
There are at most 3 bananas in the picture.

- True
- False

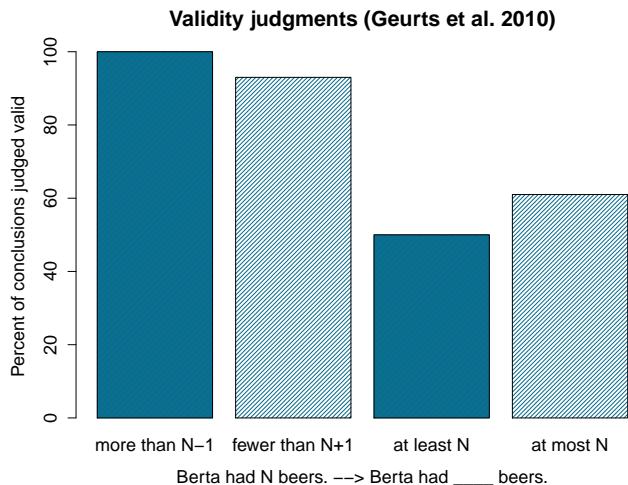
Experiment 1: Design

- 32 different pictures, with $N = 3, 4, 5$ or 6 objects (8 of each)
- 32 fillers
- 8 conditions (times 4 N s = 32 experimental stimuli):
 - *at most N and fewer than $N+1$* “true”
 - *at least N and more than $N-1$* “true”
 - *at most $N-1$ and fewer than N* “false”
 - *at least $N+1$ and more than N* “false”
- 8 lists rotating pictures through conditions
- 40 subjects (Amazon Mechanical Turk)

Experiment 1: Results (“true” conditions)



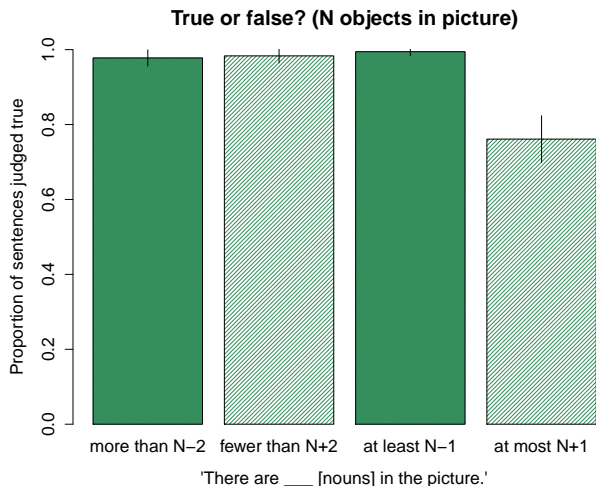
Comparison with Geurts et al. (2010)



Experiment 2: Design

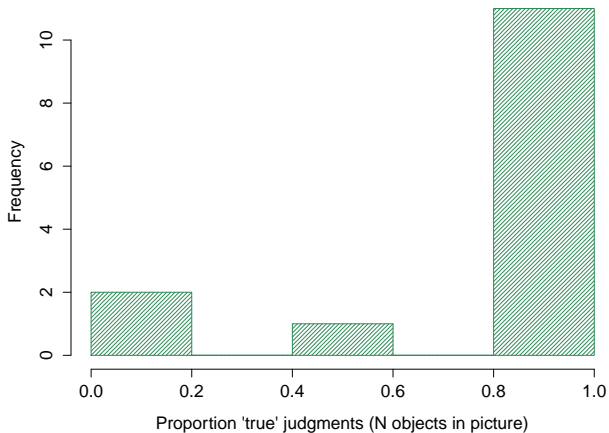
- 32 different pictures, with $N = 3, 4, 5$ or 6 objects (8 of each)
- 32 fillers
- 8 conditions (times 4 N s = 32 experimental stimuli):
 - *at most $N+1$ and fewer than $N+2$* “true”
 - *at least $N-1$ and more than $N-2$* “true”
 - *at most $N-2$ and fewer than $N-1$* “false”
 - *at least $N+2$ and more than $N+1$* “false”
- 8 lists rotating pictures through conditions
- 40 subjects (Amazon Mechanical Turk)

Experiment 2: Results (“true” conditions)



Experiment 2: Distribution for *at most N+1*

Histogram of judgments for 'at most N+1' by subject



True or false: There are at most 6 Buddas in the picture



Subject 29: "At most, there are 5." (marked it "false")

True or false: There are at most 4 candles in the picture.



Subject 31: "Technically true, but a very weird thing to say."

True or false: There are at most 6 mugs in the picture.



Subject 44: "This one is hard. I'm marking it true, but it's super-weird."

Interim conclusion

When looking at N objects:

- *at most N* is fine (Experiment 1)
- *at most $N+1$* is weirder (Experiment 2)

Experiment 3 contrasts these two conditions directly.

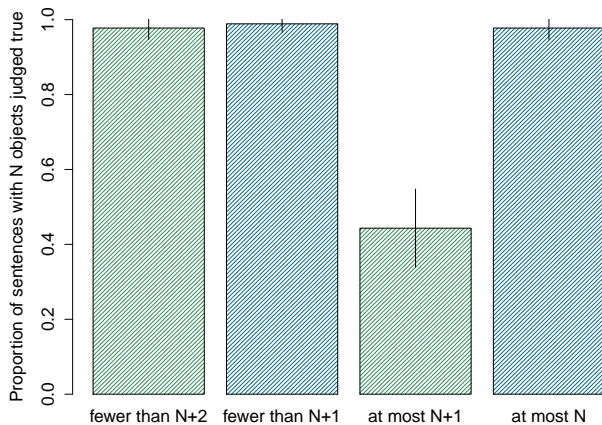
Experiment 3: Design

- 16 pictures, with $N = 3, 4, 5$ or 6 objects (4 of each)
- 32 fillers
- 4 conditions (times 4 N s = 16 experimental stimuli):
 - *at most $N+1$ and fewer than $N+2$*
 - *at most N and fewer than $N+1$*
- 4 lists rotating pictures through conditions
- 20 subjects (Amazon Mechanical Turk)

“true”

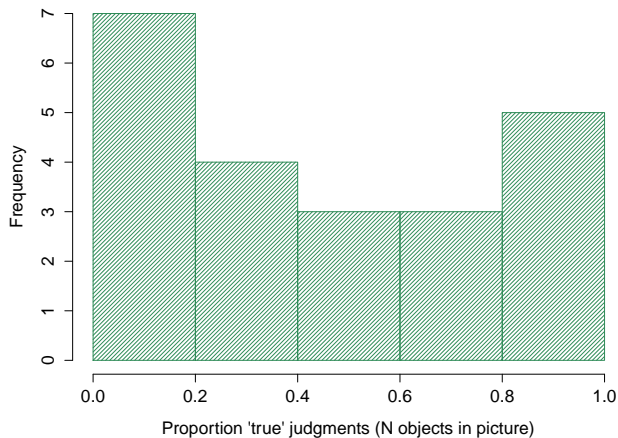
“true”

Experiment 3: Results



Experiment 3: Distribution for *at most N+1*

Histogram of judgments for 'at most N+1' by subject



Key contrast



There are at most 4 butterflies in the picture.

97%

There are at most 5 butterflies in the picture.

44%

Why this contrast?

According to Coppock and Brochhagen, i.a.:

- *There are at most 4 butterflies* means: “There are no more than 4 butterflies”, and brings attention to the possibility that there are fewer.

Why this contrast?

According to Coppock and Brochhagen, i.a.:

- *There are at most 4 butterflies* means: “There are no more than 4 butterflies”, and brings attention to the possibility that there are fewer.
- *There are at most 5 butterflies* means: “There are no more than 5 butterflies”, and brings attention to the possibility that there are fewer.

Why this contrast?

According to Coppock and Brochhagen, i.a.:

- *There are at most 4 butterflies* means: “There are no more than 4 butterflies”, and brings attention to the possibility that there are fewer.
- *There are at most 5 butterflies* means: “There are no more than 5 butterflies”, and brings attention to the possibility that there are fewer.

Both sentences are true under this analysis, and just because you bring attention to other possibilities doesn't mean you're false.

Why this contrast?

According to Coppock and Brochhagen, i.a.:

- *There are at most 4 butterflies* means: “There are no more than 4 butterflies”, and brings attention to the possibility that there are fewer.
- *There are at most 5 butterflies* means: “There are no more than 5 butterflies”, and brings attention to the possibility that there are fewer.

Both sentences are true under this analysis, and just because you bring attention to other possibilities doesn't mean you're false.

There are at least 4 butterflies brings attention to the possibility that there are 5 butterflies, but people still judge it to be true.

Analysis: Highlighting

“fewer than 6”



“fewer than 5”



“at most 5”



“at most 4”



A New Gricean Maxim

Maxim of Depictive Sincerity

If a sentence highlights a possibility, then the speaker considers it reasonably likely.

(Coppock & Brochhagen 2013)

This is a strong pragmatic requirement, so strong that it can lead people to judge true sentences as false.

Complication: Disjunctions

What is highlighted \approx what is explicitly mentioned.

Assuming disjunctions highlight both disjuncts, if one disjunct is clearly false,

N or $N+1$

should be on a par with

at most $N+1$

True or false: There are 4, 5 or 6 trumpets in the picture.



Subject 41: "I don't know what to do here - i'm NOT uncertain about the number of trumpets!" (but 94% "true", similar for other disjunctions)

True or false: There are 1 or 2 pianos in the picture.



Subject 44: “The logician in me says true, but the English speaker says false.”

True or false: There are 4 to 6 potatoes in the picture.



Subject 29: “Not really comfortable answering true or false with these range questions, but I’ll go with false again, because of the inclusion of false possibilities.”

Is the prediction supported?

Answer is mixed:

- Acceptance rates were relatively high (90% range)
- But disjunctions were the only types of items besides *at most N+1* cases for which participants volunteered a comment expressing discomfort related to the number of items.

Maxim of Depictive Inclusion

Suggestion:

- Disjunctions violate the Maxim of Depictive Sincerity.
- Depictive Sincerity violations make people uncomfortable, hence the comments, but do not on their own cause people to judge sentences as false.
- The *at most $N+1$* cases are even worse because they fail to highlight in addition the true possibility.

Maxim of Depictive Inclusion

If a sentence highlights one possibility and not another, then the speaker considers the first possibility more likely.

Outline

- 1 Introduction
- 2 Theories
- 3 Truth value judgments
- 4 Validity judgments**

Methodological conclusions from C&B

Coppock & Brochhagen (2013a) say:

- 1 Picture scenario true/false judgement tasks can cut through certain types of pragmatic infelicity that complicate the interpretation of inference judgments. (Exp. 1)
- 2 But even such true/false tasks are not impervious to particularly strong pragmatic requirements. (Exps. 2&3)

Too hasty?

There were some differences between Geurts et al.'s paradigm and C&B's, besides truth judgments vs. validity judgments:

- English vs. Dutch speakers.
- 'there are 3...' vs. 'Berta drank 3...'
- Mechanical Turk vs. paper-and-pencil
- pictures vs. no pictures

What if we eliminate these differences?

Example stimulus for validity judgment experiment

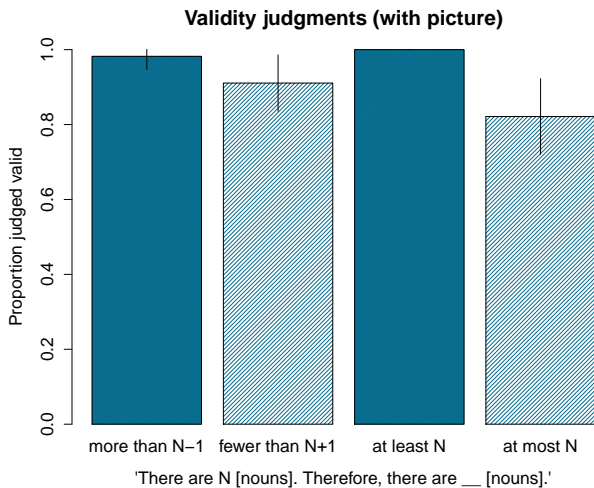
Question 1/64



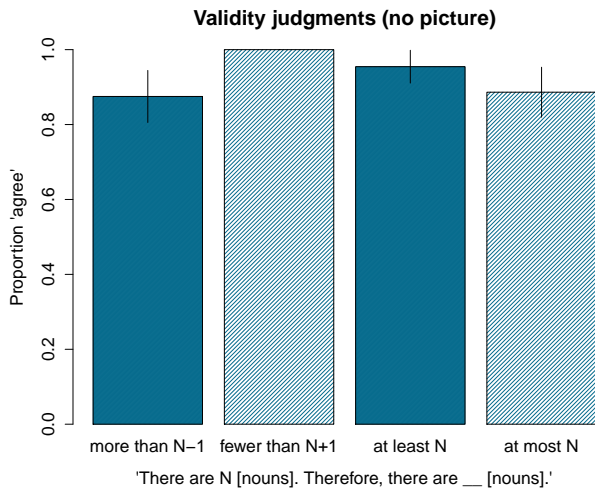
There are 3 penguins. Therefore there are fewer than 5 penguins.

- Agree
- Disagree

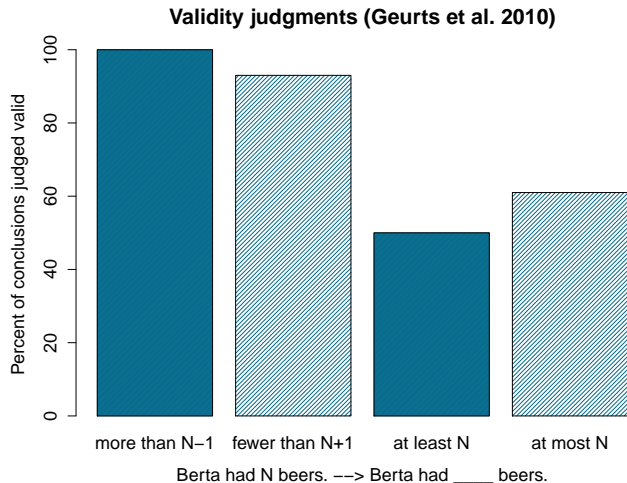
Experiment 1 with validity judgments (+ picture)



Experiment 1 with validity judgments (– picture)



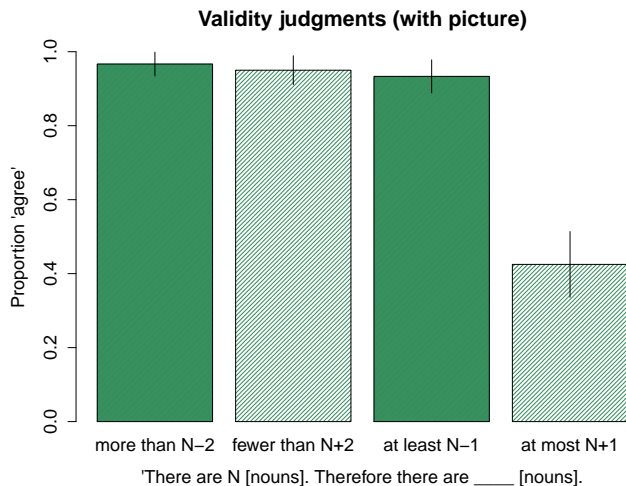
Comparison with Geurts et al. (2010)



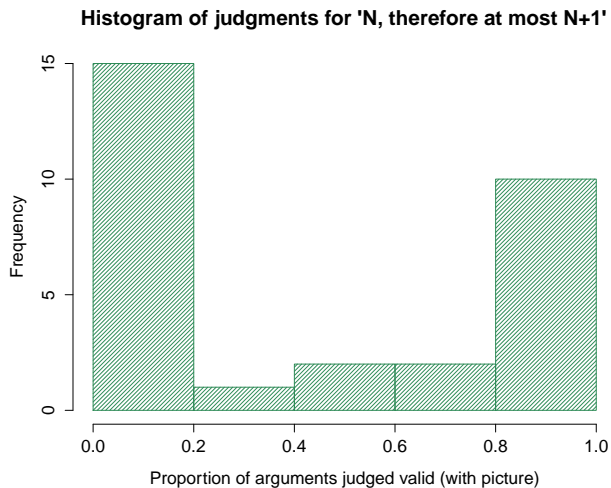
Why the failure to replicate?

- Phrasing? Spychalska (2013) has replicated other validity judgment results of Geurts et al.'s, and used slightly different phrasing.
- Different kinds of sentences ('there are...' vs. 'Berta drank...')
- Presence of other experimental items?
- Different sample of people?

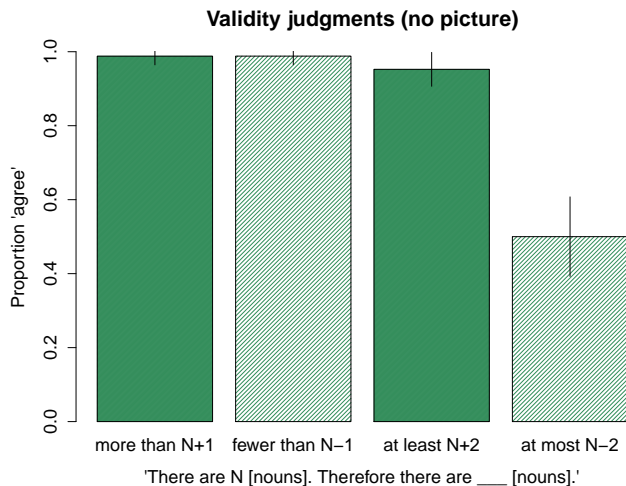
Experiment 2 with validity judgments (+ picture)



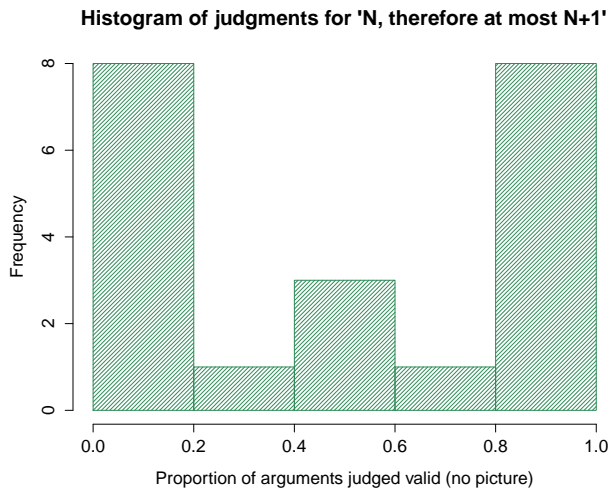
Distribution for *at most N+1* (validity + picture)



Experiment 2 with validity judgments (– picture)



Distribution for *at most N+1* (validity – picture)



Methodological conclusions

- Overall, the truth value judgment task tends to produce more categorical response patterns.
- Validity judgment tasks may be more sensitive to ignorance implicatures than truth value judgment tasks under some conditions.
- However, validity judgments do not robustly pick up on ignorance implicatures, so they cannot be relied upon for that.
- Validity judgments and truth value judgments are both sensitive to depictive sincerity/inclusion implicatures.

- Biezma, María. 2013. Only one *at least*: Refining the role of discourse in building alternatives. In *University of Pennsylvania Working Papers in Linguistics*, vol. 19 1, 11–19. Penn Linguistics Club.
- Büring, Daniel. 2008. The least *at least* can do. In Charles B. Chang & Hannah J. Haynie (eds.), *26th West Coast Conference on Formal Linguistics*, 114–120. Somerville, MA: Cascadilla Press.
- Cohen, Ariel & Manfred Krifka. 2011. Superlative quantifiers as modifiers of meta-speech acts. In Barbara H. Partee, Michael Glanzberg & Jurgis Skilters (eds.), *The Baltic International Yearbook of Cognition, Logic and Communication*, vol. 6, 1–56. New Prairie Press.
- Coppock, Elizabeth & Thomas Brochhagen. 2013a. Diagnosing truth, interactive sincerity, and depictive sincerity. In *Proceedings of SALT 23*, eLanguage.
- Coppock, Elizabeth & Thomas Brochhagen. 2013b. Raising and resolving issues with scalar modifiers. *Semantics & Pragmatics* 6(3). 1–57.
- Cummins, Chris & Napoleon Katsos. 2010. Comparative and superlative quantifiers: Pragmatic effects of comparison type. *Journal of Semantics* 27. 271–305.
- Geurts, Bart, Napoleon Katsos, Chris Cummins, Jonas Moons & Leo Noordman. 2010. Scalar quantifiers: Logic, acquisition, and processing. *Language and Cognitive Processes* 25.
- Geurts, Bart & Rick Nouwen. 2007. *At least et al.*: The semantics of scalar modifiers. *Language* 83. 533–559.

Kaplan, David. 1999. The meaning of *ouch* and *oops*. Lecture presented at the University of California at Berkeley.

Spychalska, Maria. 2013. Pragmatic effects in processing superlative and comparative quantifiers: epistemic-algorithmic approach. Slide presentation (available on author's website).