Adults' and Children's Comprehension of Disjunction in a Guessing Game: The role of measurement

Masoud Jasbi[1] & Michael C. Frank[2]

[1] Harvard University

[2] Stanford University

Author Note

All the experimental materials, data, randomization code, and analysis code for the studies reported in this paper are available in the following online repository: https://github.com/jasbi/jasbi_dissertation_LearningDisjunction. The repository also includes instructions for reproducing this research.

Correspondence concerning this article should be addressed to Masoud Jasbi, Postal address. E-mail: masoud_jasbi@fas.harvard.edu

Abstract

Previous research suggests that adults and children might differ in their interpretation of linguistic disjunction in two ways. First, children might interpret *or* as inclusive disjunction when adults interpret it as exclusive (Crain, 2012). Second, unlike adults, children might interpret a disjunction as logical conjunction (Singh, Wexler, Astle-Rahim, Kamawar, & Fox, 2016; Tieu et al., 2016). Here, we present three studies that assess adults and children's understanding of *and* and *or* using three different measures: binary forced-choice judgments, ternary forced-choice judgments, and free-form verbal feedback. Issues of measurement. Implications for pragmatic development.

Keywords: conjunction, disjunction, implicatures, semantics, pragmatics, logical connectives, language, acquisition, development, children

Word count: X

Adults' and Children's Comprehension of Disjunction in a Guessing Game: The role of measurement

Introduction

Previous research has suggested that adults and children might differ in their interpretation of *or* in two ways. First, children might interpret *or* as inclusive disjunction when adults interpret it as exclusive (Crain, 2012). Second, unlike adults, children might interpret *or* as logical conjunction, akin to *and* (Singh et al., 2016; Tieu et al., 2016). Here, we present three studies that assess adults and children's understanding of *and* and *or* in a guessing game paradigm. These studies show that four-year-olds' interpretation of conjunction and disjunction may not be as different from adults as previously supposed.

Study 1 tested adults' interpretations of logical connectives in the context of a guessing game using Two and Three-Alternative Forced Choice judgment tasks (2AFC and 3AFC). The results showed that adults interpret *and* and *or* differently. They interpreted *and* as conjunction and *or* as inclusive disjunction. However, in the task with three alternatives (3AFC) adults did not consider a disjunction felicitous when both disjuncts were true. Comparing the 2AFC and 3AFC results, we find that the felicity of disjunctive statements is sensitive to the measurement. 2AFC task systematically underestimated judgments of felicity and better approximated truth judgments compared to the 3AFC task. This finding is intuitive given that more options provide a better opportunity to express nuances of linguistic interpretation.

Study 2 investigated children's judgments in the same guessing game as study 1 using a 3AFC task. We used three alternatives to give children a better chance of expressing their pragmatic knowledge and judgments of felicity (Katsos & Bishop, 2011). The study also analyzed and categorized children's open-ended spontaneous feedback to the guesser. Both the 3AFC judgments and the categories of open-ended responses showed that four-year-olds differentiated *or* from *and*. While children's judgments in the 3AFC task showed no sign of infelicity for disjunctive guesses when both disjuncts were true, their open-ended feedback

showed that children find such guesses infelicitous. In their open-ended feedback, children's comments showed that use of a conjunction in such cases would be more appropriate.

Study 3 used the same paradigm as study 2, but focused on replicating children's open-ended responses and contrasting them with the results of a 2AFC task. As in study 2, both truth judgments and open-ended feedback showed that children differentiated *or* from *and*. The 2AFC task showed no evidence that children find disjunctions with true disjuncts infelicitous. However, children's judgments did not differ significantly from those of adults in the 2AFC task of study 1. As in study 2, children's open-ended feedback suggested that when both disjuncts are true, children find a disjunctive statement infelicitous and the conjunctive alternative more appropriate. Overall, the results of study 2 and 3 show that forced-choice judgement tasks underestimate children's pragmatic competence. Therefore, using open-ended elicitation and analysis of children's feedback **along with** forced choice judgment tasks may provide a better understanding of children's true semantic and pragmatic knowledge.

The studies reported here build on previous studies, and fill two gaps in the literature as well. First, most previous research focused on children's interpretation of *or* in complex sentences – for example with other logical words such as quantifiers *every* and *none*. Here, we test children and adults' understanding of *and* and *or* in simple existential sentences like "*There is a cat or a dog.*" To my knowledge, only Braine and Rumain (1981) used simple existential constructions before, but their experimental paradigm was relatively more complex than the paradigm used here. As discussed before, simplifying the paradigm is an important step in reducing conjunctive interpretations that arise due to non-linguistic strategies. Second, most previous research tested children and adults using 2AFC truth value judgment tasks (Crain & Thornton, 1998). Here, we report adults and children's judgments on both 2AFC and 3AFC tasks. We also use children's open-ended spontaneous feedback to develop relevant analytical response categories and we replicate the findings in a following pre-registered study. Katsos & Bishop (2011) argued that 3AFC judgment tasks are better

suited for assessing children's pragmatic competence. We present results that suggest even a 3AFC task can underestimate children's pragmatic knowledge and that children's spontaneous and open-ended elicited responses provide valuable insights not available in forced choice judgments.

Previous Research

Research on children's comprehension of logical connectives such as *and* and *or* divides into two periods. The first period (1960s-80s) was inspired by Piaget's developmental theory (Inhelder & Piaget, 1958). Researchers in this period sought to discover the development of basic logical concepts such as negation, conjunction, and disjunction. Following Inhelder and Piaget (1958), they predicted that children first form concrete concepts for conjunction and disjunction between the ages of 7-11 years (concrete operational stage) and only after 11 (formal operational stage) do they develop an abstract and logical understanding of these words. While later research in this period rejected this timeline, it confirmed the idea that a logical (inclusive) understanding of disjunction develops late. The second period (since late 90s) is inspired by Grice's theory of meaning, specifically his distinction between semantics and pragmatics. Researchers in this period argue that previous studies conflated semantic and pragmatic knowledge and used methods that vastly underestimated children' semantic competence. By controlling for the role of pramgatics and focusing on children's truth judgments, they show that children have early and adult-like semantics for logical words such as *or*. Based on these results, they argue that the understanding of logical concepts and their role in language is likely innate (Crain & Khlentzos, 2008, 2010). In what follows, I review the highlights of these two traditions and end with a note on how the research presented here contributes to this vast literature.

To examine the Piagetian theory, Nitta and Nagano (1966) tested 679 Japanese students (grades K, 2, 4, 6, and 8) and Neimark and Slotnick (1970) conducted a similar study on 455 English-speaking children in grades 3-8 and 58 college students. Participants

were tested on negation, conjunction, and disjunction. Each question provided six response options; for example a fish, a bird, and a flower, each with a white and a black version. Participants were asked to "circle all the items" described by statements such as: "flower", "not bird", "bird and flower", "bird or flower", "black and bird", "black or flower", etc. These studies concluded that the majority of the participants understood negation and conjunction, but only college students correctly answered statements containing a disjunction. They reported that participants made two types of errors. First across all ages, some participants interpreted disjunction as conjunction. For example they circled black birds when the instruction said "black or bird". Second, some selected only one of the two categories. Based on these results Neimark (1970) concluded that a "correct" (i.e. inclusive) understanding of disjunction only develops in the high school years and depends on the attainment of formal operations as defined in the Piagetian theory.

Paris (1973) used a similar in-classroom setup to test children's comprehension of connectives in Grades 2, 5, 8, 11, and college. Two hundred participants (40 per grade) were asked to judge the truth of sentences with the connectives *and*, *or*, and *either-or*. The experimenter showed participants slides of pictures, for example a bird in a nest, with descriptions such as "the bird is in the nest or the shoe is on the foot." The participants were asked to judge the statement as true or false. Paris found that statements with *and* were almost always judged correctly, but this was not the case with disjunction. First, he reported that older participants produced more errors when both disjuncts were true, presumably because they interpreted disjunctions as exclusive and not inclusive. Second, the majority of younger children, and even around a fifth of college students considered a disjunction false when only one of the disjuncts was true. The combination of these two trends suggested that initially, children did not differentiate *or* from *and*, interpreting both as conjunction. Finally, Paris also found that there were fewer errors with *either-or* statements compared to *or* statements. He suggested that the word *either* could provide further cue on how disjunction should be interpreted. Paris (1973) attributed the conjunctive interpretations of *or* to

children applying non-linguistic strategies when an utterance is hard to interpret (See Clark, 1973 for a discussion of nonlinguistic strategies in child language acquisition). He suggested that children in his task were "comparing visual and auditory information with little regard for the implied logical relationship in the verbal description." In other words, children responded with "true" if the individual disjuncts matched the pictures and false otherwise. Such a non-linguistic strategy would yield correct answers for conjunction but incorrect (conjunctive) answers for disjunction. This explains why conjunctive readings reduce with age and why using the word *either* helps reduce conjunctive interpretations further.

It was understood that the in-class tests were not suitable for testing participants' linguistic competence and certainly not suitable for younger children. Therefore, Johansson and Sjolin (1975) set out to examine the interpretaiton of disjunction in a simpler Give-item task. They tested preschool Swedish-speaking children's comprehension of conjunction and disjunction in present tense sentences (e.g. "Richard wants to drink lemonade or milk. Show me what he drank!") and imperative sentences (e.g. "Put up the car or the doll!"). They reported that starting (at least) at age four, children interpreted the Swedish equivalents of *and* and *or* as conjunction and exclusive disjunction. They argued that the linguistic *and* and *or* should be kept separate from the logical notions of conjunction and (inclusive) disjunction. While linguistic understanding of *and* and *or* develops early and in preschool years (as conjunction and exclusive disjunction), the logical understanding of them develops late.

Braine and Rumain (1981) is the only study that tested the same participants with both Give-item and Truth Value Judgment Tasks. They tested 22 children in each of the age groups 5-6, 7-8, and 9-10 years, as well as 22 adults. In the give-item task, 14 wooden blocks with varying shapes, colors, and sizes were used (a replication of Suppes & Feldman, 1969). Experimenters asked participants the following: 1) "Give me all the green things or give me all the round things" and 2) "Give me all those things that are either blue or round." They reported that for both commands and in both children and adults, the most likely response was to give all the objects that had only one of the properties. They considered these results

as evidence for a "choose-one" (i.e. exclusive) interpretation of disjunction in the context of imperatives.

In the truth value judgment task, a puppet described the contents of four boxes that each contained four animal toys. For example, the puppet said "Either there is a horse or a duck in the box." The first box had both animals, the second had only a horse, the third only a duck, and the last had neither. Participants were asked if the puppet was right. The results showed that adults were split between an inclusive and an exclusive interpretation of disjunction. The 7-8 and 9-10 year-olds were more likely to consider the disjunction as inclusive. However, the youngest group (5-6 years old) was most likely to interpret a disjunction similar to a conjunction: they said the puppet was right when both animals were in the box and not right or partly right if only one of the animals was in the box. Following Paris (1973), Braine and Rumain (1981) argued that younger chidren do not take the contribution of the connective *or* into account. Instead, they use a non-linguistic strategy in which the disjunction is right if both propositions are true, partly right if only one is true, and wrong if neither is true.

Braine and Rumain (1981) concluded that children's ability to interpret a disjunction in a command develops earlier than their ability to judge truth values. It is important to note that in Braine and Rumain (1981)'s truth value judgment task, the puppet uses a disjunction even though the content of the box was known to both the puppet and the participant (lack of ignorance). Such uses of disjunction are infelicitous. More generally, a disjunction such as "A or B" is infelicitious when discourse participants already know which proposition is true. Later truth value judgment studies such as Chierchia, Crain, Guasti, and Thornton (1998) controlled for this effect of disjunction by making the puppet utter disjunction as a prediction of an unknown event, and let participants judge the prediction after they see the outcome of the event.

Chierchia et al. (1998) kicked off the second period of inquiry into children's comprehension of disjunciton. Following Grice (1989), they differentiated between semantic

knowledge, which includes the knowledge of truth values, and pragmatic knowledge, which includes the knowldge that conversational contributions ought to be truthful, informative, relevant, and concise. They contended that interpreting logical connectives involves a semantic and a pragmatic component, and that the semantics of logical connectives cannot be assessed if the role of pragmatics is not controlled for. More specifically, they argued that felicitous use of a disjunction requires: (i) a set of alternatives (ii) evidence that one of them holds (iiia) evdience that not all of them hold, or (iiib) uncertainty as to whether all of them hold. While the semantics of *or* is inclusive, a variety of factors including pragmatic reasoning can provide evidence that not all alternatives hold. For example, we may reason that given speaker's knowledge of the situation, she could have used the connective *and* if all alternative were true. Therefore, to understand the semantic contribution of disjunction, we should test participants in contexts which are stripped from pragmatic factors that contribute to exclusivity.

They tested 23 English-speaking and 10 Italian-speaking children in two conditions: description mode and prediction mode. In both conditions, a troll considered whether to eat a hamburger, a piece of pizza, or an ice-cream for lunch and went ahead to eat a piece of pizza and an ice-cream but not a hamburger. In description mode, Kermit described what happened as "A troll ate a piece of pizza or an ice cream" while in prediction mode, Kermit used the same sentence as a prediction before the troll eats his lunch. They reported that in the description mode, children accepted Kermit's statement when both disjuncts were true less than one-third of the time. However, in prediction mode, they accepted such sentences 100% of the time. They argued that when we control for the effect of pragmatics on interpretation, children understand disjunction as inclusive, and conform to the semantics of disjunciton in classical logic.

Following Chierchia et al. (1998), several studies have argued that preschool children's knowledge of disjunction conforms to the predictions of classical logic and formal semantics in environments as varied as negative sentences (Crain, Gualmini, & Meroni, 2000),

conditional sentences (Gualmini, Crain, & Meroni, 2000), restriction and nuclear scope of the universal quantifeir *every* (Chierchia, Crain, Guasti, Gualmini, & Meroni, 2001; Chierchia et al., 2004), nuclear scope of the negative quantifier *none* (Gualmini & Crain, 2002), restriction and nuclear scope of *not every* (Notley, Thornton, & Crain, 2012), and prepositional phrases headed by *before* (Notley, Zhou, Jensen, & Crain, 2012), as well as similar environments in other languages such as Mandarin Chinese and Japanese (Goro & Akiba, 2004; Su, 2014; Su & Crain, 2013). These studies also commonly reported that in linguistic environments where adults consider a disjunction exclusive, children are more likely to consider it inclusive. Since under the Gricean account, exclusive interpretation of disjunction is the result of pragmatic (scalar) implicaatures, these findings are considered as further evidence for the hypothesis that young children do not compute implicatures at the rate that adults do (**???**; Noveck, 2001).

It is important to note that all the studies mentioned above in the Gricean period use the Truth Value Judgment Task as specified in (Crain & Thornton, 1998). As mentioned earlier, Braine and Rumain (1981) found that the same children were more likely to interpret a disjunction as exclusvive in a give-item task and inclusive/conjunctive in a truth value judgment task. Therefore, it is possible that truth value judgment tasks are simply not suitable for capturing children' knowledge of exclusivity implicatures. Furthermore, several studies listed above test children's knowledge of disjunction in environments that largely collapse the distinction between *and* and *or*. For example, in the restriction of *every*, a conjunction and a disjunction can result in the same interpretation (e.g. *Every man or woman is happy* vs. *Every man and woman is happy*). Therefore, successful interpretation in these studies can also be achieved by applying the nonlinguistic strategies that result in conjunctive interpretations, as discussed by the early studies in the Piagetian period.

More recently, some studies have revived the earlier findings that preschool children may interpret disjunction as conjunction. Singh et al. (2016) tested 56 English-speaking children (M=4;11, 3;9-6;4) and 26 adults in a truth value judgment task. The experiment

involved four pictures: a boy holding a banana, a boy holding an apple and a banana, three boys holding either an apple or a banana, and three boys holding both apples and bananas. In each trial, participants saw one of the pictures and a puppet described the pictures with four possible utterances: "The/every boy is holding an apple or/and a banana." Participants were asked: "Was [the puppet] right or wrong about this picture?" They found that children were more likely to say the puppet was right when both disjuncts were true than when only one was. They concluded that "many preschool children - the majority in [the study's] sample - understand disjunctive sentences ... as if they were conjunctions."

Tieu et al. (2016) also found evidence for conjunctive interpretations of disjunction in preschool children. They tested 28 French-speaking children (3;7-6;6, M=4;5) and 18 Japanese-speaking children (4;7-6;6, M=5;5) as well as 20 French-speaking and 21 Japanese-speaking adults. They used the "prediction mode" of the Truth Value Judgment Task, in which the puppet provides a prediction or guess, an event occurs, and participants are asked if the prediction was right. For example, there was a chicken on the screen, next to a toy bus and a toy plane. The puppet appeared on the screen and predicted that "the chicken pushed the bus or the plane." Then the chicken pushed either one or both of the objects. Participants stamped on a happy face or a sad face to show whether the puppet's guess was right or wrong. Like Singh et al. (2016), they reported that unlike adults, children were more likely to consider the disjunctive guess right when both disjuncts were true, rather than only one. They concluded that children - the majority of them in their sample - interpreted disjunction as conjunction.

However, a recent replication of Tieu et al. (2016) by Skordos, Feiman, Bale, and Barner (2018) suggests that the high rate of conjunctive interpretations were most likely due to experimental design. They tested 126 preschoolers in three conditions: replication (N=43, 4;0-5;9, M=5;0), modified script (N=41, 4;0-5;10, M=5;0), and three-alternatives (N=42, 4;0-5;11, M=5;0). The first condition was a direct replication of Tieu et al. (2016). The second, modified script, removed some experimenter comments right after the puppet's guess

that could potentially confuse children. The comments were: "Look! The chicken pushed that! She didn't want to break that one. So she didn't touch it. So was [the puppet] right?" The third condition, three-alternatives, was similar to modified-script but provided three objects; for example a plane, a bus, and a bicycle. The reasoning was that if there are only two alternatives, a disjunction is trivially true, and consequently children may consider that unacceptable. The results replicated Tieu et al. (2016)'s findings in the replication condition, but showed that conjunctive interpretations of disjunction disappeared almost completely in the third condition with three alternatives. Skordos et al. (2018) concluded that children's conjunctive interpretations are most likely due to non-linguistic strategies applied when they are uncertain about some aspect of the experimental task. This conclusion is similar to that of earlier studies on conjunctive interpretations of disjunction in the 70s and 80s.

To summarize, previous studies show that the design of experimental tasks can have a big impact on our conclusions regarding children's comprehension of disjunction. Early in-class tasks suggested that even high-schoolers do not interpret a disjunction correctly and confuse it with *and*. Improving on task design, Braine and Rumain (1981) argued that this is only the case in preschool children. They also showed that the same children can have different interpretations of disjunction in different tasks: in a give-item task they interpret it as exclusive while in a truth-value judgment task they interpret it as conjunctive or inclusive. Using various versions of the truth value judgment task, research in the Gricean tradition has argued that preschool children understand the semantics of disjunciton and interpret it as inclusive. However, this line of research has largely suggested that children are insensitive to the exclusivity implicature of disjunciton. While some recent studies have argued that preschool children may interprete disjunction as conjunctive, a replication study has argued that conjunctive interpretations were largely due to task demands.

Here we improve on previous studies by first controling for various factors that had proven problematic for previous studies, and second investigating the role of measurement in preschool children's interpretation of disjunction. As explained above, previous research has

shown that in studying children's interpretation of disjunction, it is important to control for the following factors:

1. complexity of the linguistic stimuli,
2. complexity of the task,
3. ignorance of the speaker with respect to the truth of the disjunts,
4. interpretation of the conjunction word (e.g. *and*) in the same task
5. interpretation of adults in the same task.
6. Discernibility of conjunctive and disjunctive interpretations in the task

Some previous studies used complex linguistic stimuli or relatively complex designs that may have increased the application of non-linguistic strategies. Some studies violated "speaker ignorance"; i.e. had the speaker utter the disjunction when the truth of the propositions were known to the speaker. Some studies did not use the conjunction word (e.g. *and*) in control trials, or did not use adults as control participants. Finally, some studies tested the disjunction word in linguistic environments that collapse interpretive differences between the conjunction and disjunction words. The experimental paradigm reported here builds and improves on previous studies by controlling for all these factors.

In the studies reported here, we used simple existential sentences (e.g. *there is a cat or a dog*) and tested the interpretation of participants in a simple and easy to understand guessing game. The guessing game provided a context in which the speaker was ignorance with respect to to which alternatives actually hold. The game is essentially a variant of the truth value judgment task. The study used conjunction trials as well as adult participants as controls. The conjunction word *and* and the disjunction word *or* resulted in different interpretations in the task. Furthermore, we tested children's interpretations in two different ways, using forced choice tasks with 2 and 3 options, as well as free form verbal responses.

Study 1: Adult's 2AFC and 3AFC Judgments

The goal of this study was to examine adults' interpretations of *and* and *or* as a benchmark for children's interpretations. We designed the study as a guessing game. Participants saw a card, read a description, and had to evaluate the description with respect to what they saw on the card. In test trials, the descriptions contained the conjunction word *and* and the disjunction word *or*. We tested adults in both two-alternative and three-alternative forced choice tasks (2AFC and 3AFC). The results suggested that adults interpreted *and* as conjunction and *or* as inclusive disjunction. Adults also considered statements with *or* infelicitous when both disjuncts were true. The study also found that the 2AFC and 3AFC tasks registered different aspects of adult interpretations: the 2AFC task captured adult intuitions on the basic semantics of the connectives while the 3AFC task was sensitive to pragmatic infelicities as well.

Methods.

Materials and Design. We used six cards with cartoon images of a cat, a dog, and an elephant (Figure 1). There were two types of cards: cards with only one animal and cards with two animals. There were three types of guesses: simple (e.g. *There is a cat*), conjunctive (e.g. *There is a cat and a dog*), and disjunctive (e.g. *There is a cat or a dog*). In each guess, the animal labels used in the guess and the animal images on the card could have no overlap (e.g. Image: dog, Guess: *There is a cat or an elephant*), partial overlap (e.g. Image: Cat, Guess: *There is a cat or an elephant*), or total overlap (e.g. Image: cat and elephant, Guess: *There is a cat or an elephant*). Crossing the number of animals on the card, the types of guesses, and the overlap between the guess and the card yields 12 different possible trial types. We chose 8 trial types (Figure 2), to balance the number of one-animal vs. two-animal cards, simple vs. connective guesses, and expected true vs. false trials.

Participants and Procedure. We used Amazon's Mechanical Turk (MTurk) for recruitment and the online platform Qualtrics for data collection and survey design. The task took about 5 minutes on average to complete. 109 English speaking adults participated.

57 of them were assigned to a 2AFC judgment task and 52 to a 3AFC judgment task. In the 2AFC task, participants had to judge using the options "wrong" and "right". In the 3AFC task they had to choose between "wrong", "kinda right", and "right". The two conditions were otherwise identical. There are many possible labels for the middle option "kinda right", including "kinda wrong" or "neither". A later experiment, tested different intermediate labels and found that adults consider "kinda right" to be a more suitable option for capturing pragmatic infelicities (see Jasbi, Waldon, & Degen, submitted). We expect similar behavior from labels like "a bit right" and "a little right" which refer to non-maximal degrees of being "right".

The experiment had three phases: introduction, instruction, and test. In the introduction, participants saw the six cards and read that they would play a guessing game. Then a blindfolded cartoon character named Bob appeared on the screen. Participants were told that in each round of the game, they would see a card and Bob was going to guess what animal was on the card. The study emphasized that Bob could not see anything. Participants were asked to judge whether Bob's guess was right. In the instruction phase, participants saw an example trial where a card with the image of a dog was shown with the following sentence written above Bob's head: *There is a cat on the card.* All participants correctly responded with "wrong" and proceeded to the test phase.

In the test phase, participants saw one trial per trial type. Within each trial type, the specific card-guess scenario was chosen at random. The order of trial types was also randomized. At the end of the study, participants received $0.4 as compensation. Figure 3 shows an example test trial.

Table 1

Summary of study 1 methods with adults participants

Study	N	Age	Mode	Response Options
Study 1 - Part 1	57	Adults	Online (Mturk)	Wrong, Right

Study	N	Age	Mode	Response Options
Study 1 - Part 2	52	Adults	Online (Mturk)	Wrong, Kinda Right, Right

Results. In this section, we first present the results of the 2AFC and 3AFC tasks with adults. Then we discuss how these results can be interpreted with respect to the semantics and pragmatics of disjunction in the context of the guessing game.

Judgments with Two Alternatives (2AFC). Figure 4 shows the results for the adult 2AFC task. The two left columns show the simple guesses and serve as controls. The results show that if the animal mentioned in the guess was not on the card (e.g., elephant), participants judged the guess to be "wrong"; if the animal was on the card (e.g., cat), participants judged the guess to be "right". The next two columns of Figure 4 show the results for the test conditions, namely conjunction and disjunction. An *and*-guess (e.g. cat and dog) was considered "wrong" if only one of the animals was on the card, and "right" if both were. An *or*-guess (e.g. cat or dog) was "right" whether one or both animals were on the card. The patterns of "right" and "wrong" responses in the binary task match the expectations for truth and falsehood of logical conjunction and (inclusive) disjunction.

Judgments with Three Alternatives (3AFC). Figure 5 shows the results for the 3AFC judgment task. For four trial types, the results were identical to the 2AFC task. In the first and second trial types, if the animal mentioned was not on the card (e.g. elephant), participants judged the guess as "wrong", regardless of wether one animal was on the card or two. In the third trial type, if the animal mentioned (e.g. cat) was the only animal on the card, participants judged the guess as "right". Finally, if there were two animals on the card and the puppet mentioned them using *and* (e.g. cat and dog), all participants considered the guess "right".

The four remaining trial types showed different patterns of judgments than the ones in the 2AFC task. If the animal mentioned (e.g. cat) was only one of the animals on the card, participant judgments were divided between "right" and "kinda right" (See Table 2, row 1

for the statistical test). Also, most adults considered a conjunctive guess (e.g. cat and dog) "wrong", when only one of the animals was on the card (Table 2, row 2). However, some considered it "kinda right", perhaps suggesting that the intermediate option was used to express the notion of partial truth. With respect disjunctive guesses (e.g. cat or dog), if the card had only one of the animals, most adults considers the guess "right" while some considered it "kinda right" (Table 2, row 3). It is possible that the adults who considered such guesses "kinda right" were sensitive to the under-informative nature of a disjunctive guess when a simple guess like "cat" would have been more appropriate. If both animals were on the card, adults were split between "kinda right" and "right" responses (Table 2, row 4). The choice of "kinda right" over "right" in such trials can be interpreted as a sign that adults were sensitive to the infelicity of a disjunction when conjunction was more appropriate. However, the scalar reasoning with *and* and *or* is subtle and in section , we discuss the nature of this reasoning in the context of this guessing game.

Table 2

Exact One-Sided Binomial Test

Trial Type	n_{right}/n_{total}	$\hat{p}_{right}$	p_{null}	P-value	95% CI
Two Animals - Simple	32/52	0.62	0.50	0.06	0.49-1
One Animal - AND	16/52	0.69	0.50	0.00	0.57-1
One Animal - OR	19/52	0.63	0.50	0.04	0.51-1
Two Animals - OR	32/52	0.62	0.50	0.06	0.49-1

Discussion. The example sentences bellow show the common interpretations of conjunctive and disjunctive assertions (Aloni, 2016).

- Bob is sad *and* angry.

- Both are true. (Truth Conditional Meaning)

- Bob is sad *or* angry.
  - At least one of the two is true. (Truth Conditional Meaning)
  - Speaker doesn't know which is true. (Ignorance Inference)
  - At most one of the two is true. (Exclusivity Inference)

A conjunctive assertion implies that both propositions are true while a disjunctive assertion implies that at least one is true. These two inferences follow from the classical truth-conditional account of conjunction and disjunction. They constitute the semantics of *and* and *or*. However, a disjunctive assertion often has two additional inferences: an ignorance inference and an exclusivity inference. These additional inferences are often classified under pragmatic meaning. This section discusses the semantics and pragmatics of *and* and *or* in the context of the guessing game in Study 1.[1]

The Semantics of AND and OR. Let's assume that the semantics of *and* and *or* in simple declarative sentences like "there is a cat or(and) a dog" is captured by the logical operators conjunction and inclusive disjunction respectively. A conjunction is true when both conjuncts are true and false otherwise. An inclusive disjunction is true when at least one disjunct is true and false otherwise. Let's also assume a simple linking function in which false statements are judged as "wrong" and true statements as "right" (see Jasbi et al. (submitted) for a discussion of linking assumptions in this task). In the context of study 1, this purely semantic (i.e. truth-conditional) account has two main predictions: 1. Conjunctive guesses like "cat and dog" are wrong when only one of the animals is on the card. 2. Disjunctive guesses are always right because in all such trials at least one of the animals is present on the card. Figure 4 shows that in 2AFC judgments, both predictions are borne out. In other words, judgments with two alternatives seem to match the predictions of

[1] See Gutzmann (2014) for a comprehensive discussion of the definitions and boundaries of semantics and pragmatics. Here my definitions and assumptions are close to those of Gazdar (1979).

a purely semantic account of the connectives *and* and *or* with a linking function that considers "right" and "wrong" roughly as "true" and "false".

However, in the 3AFC task, judgments deviated from a purely semantic account in four trial types: 1. disjunction trials with one animal 2. disjunction trials with two animals, 3. conjunction trials with one animal, and 4. trials with simple guesses when two animals were shown on the card. Participants often used the third option "kinda right" in these trial types. Other trial types obtained identical results in 2AFC and 3AFC tasks. The comparison of forced choice judgments with two and three alternatives suggests that two alternatives better captured the truth-conditional meaning of the connectives, but underestimated adult pragmatic reasoning in the guessing game.

The Pragmatics of AND and OR. A disjunctive assertion like "cat or dog" gives rise to an ignorance inference and an exclusivity inference. The ignorance inference is the inference that the speaker does not know which disjunct actually holds. For example in figure 6, the disjunctive guess is uncertain between three outcomes: cards 1, 2, and 3. A disjunction is infelicitous when the outcome is known to discourse participants. For example, Tarski (1941) mentioned that a disjunction like "the grass is green or blue" is odd because we already know that the grass is green. The guessing game in this study controls for this ignorance effect by keeping the guesser blindfolded. Therefore, all the disjunctive guesses are evaluated in a context where participants know that the guesser is ignorant of the animals on the cards - both the number of them on the card and their identity. The exclusivity inference is the inference that only one of the disjuncts holds and **not both**. In figure 6, a disjunction like "cat or dog" only refers to cards 2 and 3 if it is accompanied by an exclusivity inference.

Since Grice (1989), this exclusive interpretation of *or* has been (at least partly) attributed to pragmatic reasoning about the speaker's connective choice. The reasoning goes like this: conversational participants are required to make their utterances as informative as possible. In the context of making predictions and guessing, a guesser is required to make

any guess as specific (i.e. informative) as possible.[2] A conjunction is more specific and informative than a disjunction (Horn, 1989). For example in Figure 6, *cat and dog* picks card 1 while *cat or dog* refers to cards 1, 2, and 3. If speakers intend to refer to card 1, they should use *and* and say *cat and dog*. If they use *or* instead of *and*, they probably do not intend to refer to card 1. Following this line of reasoning, we can exclude the possibility that a speaker intends to refer to card 1. The term "exclusivity implicature" captures this pragmatic reasoning that results in excluding the possibility of both disjuncts being true.

Our goal here is to lay out the structure of the exclusivity reasoning in the experimental setup and explain how it may be manifested in the results of the experimental studies. There are three main components to the pragmatic reasoning in the guessing game: 1. the assumptions of the game. 2. sensitivity to (under)informativity, and 3. the pragmatic reasoning about the speaker's choice of connectives. Like Katsos & Bishop (2011), we have considered "sensitivity to informativeness" as a precondition for "derivation of scalar implicatures". We begin with the assumptions of the guessing game.

- **Guessing Game Assumptions**:
  - **Ignorance**: the guesser does not know the number or identity of the animals on the card.
  - **Specificity**: the guesser is required to be as specific as possible, ideally referring to a single card.

As explained before, ignorance of the guesser was explicit and part of the instructions in the study. However, specificity was an implicit assumption[3]. All the guesses used in the experiment can pick a single card except for disjunctive ones. Conjunctive guesses like *cat*

[2]When you ask someone to predict the outcome of a coin toss, a guess like "it will be heads or tails" does not count as a felicitous guess or prediction, presumably because it is not informative, ie.e. it will always be true.

[3]Making this assumption explicit is both hard for young children and almost impossible when disjunctive guesses are used. Disjunctive guesses are always underinformative and never pick out a specific card.

and dog pick specific cards. The simple ones like *cat* can be strengthened pragmatically to mean "only a cat", and pick a specific card. However, Disjunctive ones like *cat or dog* pick two cards in their most specific (exclusive) sense. Therefore, they are always under-informative and violate the specificity assumption.

- **Sensitivity to Informativeness**: The guesser said *cat or dog* which is under-informative and picks cards 1, 2, and 3.
- **Violation Assumption**: the guesser is violating the specificity requirement.

Participants can detect the underinformativity of disjunctive guesses, notice the violation of specificity, and then decide whether they would like to tolerate this violation or punish it. It should be pointed out that it is hard to distinguish between "tolerating the specificity violation" and simply revising the specificity assumption of the game to avoid a violation. For example, participants may assume that the goal of the game is saying something true about the cards rather than being as specific as possible. In either case, the prediction is that adults who tolerate violation or revise specificity would judge disjunctive guesses as "right". However, if participants assume specificity and decide to not tolerate its violation, they will judge all disjunctive guesses to have some degree of infelicity. Since an under-informative guess is still technically correct, participants may not punish such a guess with a "wrong" response and prefer an intermediate option like "kinda right". This is what study 1 shows. With two alternatives, not many adults judge infelicity with disjunctive guesses and there are almost no "wrong" responses. With three alternatives, "kinda right" responses pop up. Adult responses are split between "kinda right" and "right".

If detecting and reacting to underinformativity is the whole story, then disjunctive guesses should show similar degrees of infelicity, regardless of how many animals there are on the card. However, the results of the 3AFC task suggest otherwise. A logistic mixed-effects model with the random intercepts and slopes for subjects and fixed effect of disjunction type found that when comparing disjunctive guesses in the 3AFC task, participants were more

likely to choose "kinda right" than "right" when both animals were on the card (β=-1.22, z=-2.25, p=0.02). In other words, participants judged further infelicity with disjunctive guesses that had both disjuncts as true. Therefore, it is possible that in some trials when both disjuncts were true, some participants went through the following pragmatic reasoning:

- **Reasoning on Alternatives**: Why did the guesser choose the under-informative connective *or* rather than the more informative *and*?
- **Resolution Assumption**: speaker is trying to be as specific as possible by resolving the issue of how many animals are on the card.
- **Exclusivity Implicature**: Given the resolution hypothesis, if the speaker had decided that two animals were on the card, they should have said *cat and dog*. They did not, so they had decided that only one animal is on the card and not both.

How does the exclusivity implicature affect participant judgments in the experimental setting? One possibility is that excluding the correct response pragmatically is treated like cases of excluding the right response semantically. For example, guessing "elephant" when there is a cat on the card. The prediction is that disjunctive trials with true disjuncts should receive "wrong" responses. However, this prediction was not borne out. Such disjunctive trials are almost never judged as "wrong".

Alternatively, it is possible that adults differentiate incorrect pragmatics from incorrect semantics (i.e. falsehood) and punish incorrect pragmatics less than incorrect semantics. This conclusion is supported by the response patterns across trial types (figure 5). Trial types that received a "wrong" response were those that were false. Pragmatically infelicitous trial types, namely simple guesses like *cat* or disjunctive guesses like *cat or dog* when both animals are on the card, receive "kinda right" responses. In other words, adults consider false utterances as "wrong" guesses but infelicitous utterances do not reach the level of being "wrong"; they are still right even though not completely right. This would explain why the

rates of infelicity (avoiding the "right" alternative) differ between 2AFC and 3AFC tasks in disjunctive trials with true disjuncts (0.18% vs. 0.62%).

Study 2: Children's 3AFC judgments and open-ended verbal feedback

The goal of this study was to examine children's interpretations of *and* and *or* in the guessing game and compare them to those of the adults. Since the 3AFC judgment task in study 1 proved better at capturing the nuances of adults' pragmatic reasoning, we decided to first test children using the 3AFC task. We also analyzed children's open-ended verbal feedback about the guesses in the experimental context. Both 3AFC judgments and the analysis of children's open-ended feedback showed that children differentiate existential sentences with *and* from those with *or*. While the 3AFC task suggested that children consider disjunctive guesses with true disjuncts as felicitous, the analysis of their verbal feedback showed otherwise. Children took issue with such guesses and corrected them, often by mentioning the stronger alternative *and*. We conclude that the 3AFC task may have underestimated children's pragmatic competence.

Table 3

Summary of Study 2 Methods

Study	N	Age	Mode	Response Option
Study 2	42	3;1-5;2 (M = 4;3)	Study Room	Circle (wrong), Little Star (little right), Big Star (right)

Methods.

Materials and Design. We used the same set of cards and linguistic stimuli as the ones in study 1. There were 8 trial types and 2 trials per trial type for a total of 16 trials. We made two changes to make the experiment more suitable for children. First, instead of the fictional character Bob, a puppet named Jazzy played the guessing game with them. Jazzy wore a sleeping mask over his eyes during the game (Figure 8). Second, a pilot study

showed that a scale with three alternatives is better understood and used by children if it is presented in the form of rewards to the puppet rather than verbal responses such as "wrong", "a little bit right", and "right", or even hand gestures such as thumbs up, middle, and down. Therefore, we placed a set of red circles, small blue stars, and big blue stars in front of the children. These tokens were used to reward the puppet after each guess. During the introduction, the experimenter explained that if the puppet is right, the child should give him a big star, if he is a little bit right, a little star, and if he is not right, a red circle.

Participants and Procedure. We recruited 42 English speaking children from the Bing Nursery School at Stanford University. Children were between 3;1 and 5;2 years old (Mean = 4;3). The experiment was carried out in a quiet room and all sessions were videotaped. There was a small table and two chairs in the room. Children sat on one side of the table and the experimenter and the puppet on the other side facing the child. The groups of circles, small stars, and big stars were placed in front of the child from left to right. A deck of six cards was in front of the experimenter. As in study 1 with adults, the children went through three phases: introduction, instruction, and test.

The goal of the introduction was for the experimenter to show the cards to the children and make sure they recognized the animals and knew their names. The experimenter showed the cards to the children and asked them to label each animal. All children recognized the animals and could label them correctly. In the instruction phase, children went through three example trials. The experimenter explained that he was going to play with the puppet first, so that the child could learn the game. He removed the six introduction cards and placed a deck of three cards face-down on the table. From top to bottom (first to last), the cards had the following images: cat, elephant, cat and dog. He put the sleeping mask on the puppet's eyes and explained that the puppet is going to guess what animal is on the cards. He then picked the first card and asked the puppet: "*What do you think is on this card?*" The puppet replied with "*There is a dog*". The experimenter showed the cat-card to the child and explained that when the puppet is "not right" he gets a circle. The pilot study had

shown that some children struggle with understanding the word "wrong", so "not right" was used instead. He then asked the child to give the puppet a circle. Rewards were collected by the experimenter and placed under the table to not distract the child. The second trial followed the same pattern except that the puppet guessed "right" and the experimenter invited the child to give the puppet a big star. In the final trial, the puppet guessed that there is a cat on the card when the card had a cat and a dog on it. The experimenter said that the puppet was "a little right" and asked the child to give him a little star.

Table 4

Instruction Trials.

Card	Guess	Reward
CAT	There is a cat!	Circle
ELEPHANT	There is an elephant!	Big Star
CAT-DOG	There is a dog!	Little Star

In the test phase, the experimenter removed the three instruction cards and placed a deck of 16 randomized cards on the table. The experimenter explained that it was the child's turn to play with the puppet. The test phase followed the pattern described in the instruction phase.

Offline Annotations. During analysis of the videos, children's linguistic feedback to the puppet after each guess was categorized into four types: 1. None, 2. Judgments, 3. Descriptions, and 4. Corrections. The first category referred to cases where children did not say anything and only rewarded the puppet. Judgments referred to linguistic feedback such as *you are right!*, *yes*, *nope*, or *you winned*. Such feedback only expressed judgments and complemented the rewards. Descriptions were cases that the child simply mentioned what was on the card: *cat!*, *dog and elephant!*, *There is a cat and a dog!* etc. Finally, corrections referred to feedback that provided "focused elements" that acted like corrections to what the puppet had said. Examples include: *Just a cat!*, *Both!*, *The two are!*, *Only cat*, *cat AND dog*

(with emphasis placed on *and*). In trials where the child provided both judgments as well as descriptions or corrections, we placed the feedback into the more informative categories, namely description or correction.

Results. Figure 9 shows the results for children's 3AFC judgments. Starting from the left column, if the mentioned animal was not on the card (e.g. elephant), children judged the guess as "wrong". If the animal mentioned (e.g. cat) was the only animal on the card, children judged the guess to be "right". Here we ignore the results for trial types in which the animal mentioned was one of the animals on the card. The reason is that such trials were used in the instruction phase to introduce the "little bit right" guesses, and the results are potentially biased by the instructions.

In conjunctive guesses (e.g. *cat and dog*), when only one of the animals mentioned was on the card, children judged the guess as "wrong" or "a little bit right". However, if both animals were on the card, they judged the conjunctive guess as "right". In disjunctive guesses (e.g. *cat or dog*), when only one of the animals mentioned was on the card, children considered the guess "right" or "kinda right". If both animals were on the card, the disjunctive guess was considered "right".

The comparison of conjunction and disjunction trials (last two columns of figure 9) shows that overall, children distinguished between *and* and *or* when one animal was on the card. Given that the one-animal conjunction trials are false but the one-animal disjunction trials are true, the difference in response patterns may suggest that children understood the truth-conditional differences between *and* and *or*. The truth judgments did not provide evidence that children differentiated *and* and *or* when two animals were on the card. Since in the majority of examples with *or* and two animals, children responded with "right", it is possible to conclude from the 3AFC judgment data that children did not generate exclusivity inferences in this task.

Figure 10 compares the results for children and adults' 3AFC judgments in the conjunction and disjunction trials. The major difference between adults and children's

responses was disjunctive trials with two animals on the card. Most children considered such trials as "right" while adults considered them as "kinda right". In the next section, we use Bayesian regression modeling to compare adults' and children's three-alternative responses more systematically.

Analysis and Statistical Modeling. We used the R package RStan for Bayesian statistical modeling to fit separate ordinal mixed-effects logistic models for the children's and adults' judgments. The response variable had three ordered levels: *wrong*, *kinda right*, and *right*. The trial types *One-Animal-OR*, *Two-Animals-OR*, *One-Animal-AND* constituted the (dummy-coded) fixed effects of the model with *Two-Animals-AND* set as the intercept. The model also included by-subject random intercepts. The priors over trial types and the random intercepts were set to $\mathcal{N}(0, 10)$. We also included parameters C_1 and C_2, the two cutpoints delimiting the logistic for 1) *wrong* and *kinda right* and 2) *kinda right* and *right* responses, drawn with the prior $\mathcal{N}(0, 1)$.[4] All four chains converged after 3000 samples (with a burn-in period of 1500 samples).

We made inferences based on the highest-posterior density (HPD) intervals for the coefficients estimated from each model. Because predictors are dummy-coded, it's possible to examine contrasts of interest by computing the difference between coefficients for pairs of conditions we wish to contrast. In naming the coefficients like b(OR,2), OR/AND represents the connective used and the number represents the number of animals on the card. Figure 11 shows the contrasts of interest: *b(OR, 2)-b(OR,1)* represents the difference between the estimated coefficients for the disjunction trials with two animal on the card and those with only one; *b(OR, 2)* represents the difference between the estimated coefficients for the conjunction trials with two animals and the disjunction trials with two animals; and so on.

Overall, adults' and children's estimated coefficients are similar in sign to one another, though adults' are more extreme. In the conjunction trials (*b(AND, 2)-b(AND,1)*), children and adults showed a strong preference for the cards with two animals rather than one. At

[4] We used a tight prior in this case to decrease posterior correlations between cutpoints and intercept.

the same time, given two animals on the card, children and adults showed a preference for *and* rather than *or* (*b(OR, 2)*). However, with only one animal on the card, children and adults preferred a disjunctive guess (*b(OR, 1)-b(AND,1)*). These results are compatible with the truth conditions of conjunction and disjunction.

The main difference between adults and children shows up in the contrast between the disjunctive trial types: two animals vs. only one (*b(OR, 2)-b(OR,1)*). On average, children rated disjunction trials with two animals higher than those with only one. Adults on the other hand showed the opposite pattern: they rated disjunction trials with two animals lower. This pattern is compatible with current accounts of pragmatic development that suggest an absence of implicatures in children's interpretations. The idea is that while adults strengthen the disjunctive guess *cat or dog* to "cat or dog but not both", children simply interpret it as "cat or dog or both". Adults are therefore going to rate trials with both disjuncts true lower.

The slight preference children show for cards with two animals when the guess is disjunctive is also compatible with the account proposed by Singh et al. (2016) and Tieu et al. (2016). However, the effect seems much smaller here than was reported in their studies. The comparison with conjunction trials makes it clear that overall, children are not interpreting *or* as having a conjunctive meaning. The effect in this study can be more accurately described as a preference in judgment for both disjuncts being true rather than a conjunctive interpretation of disjunction. The results from children's spontaneous linguistic feedback make it less likely that children interpretive *or* as a conjunction. We will discuss this issue further in section .

Table 5

Definitions and Examples for the Feedback Categories.

Category	Definition	Examples
None	no feedback provided to the puppet, only reward	

Category	Definition	Examples
Judgment	the child said yes/no, you are right, etc.	"No!" , "You are right Jazzy!"
Description	mentioned the animal(s) on the card	"elephant", "cat and dog"
Correction	used focus particles like *only/just*, emphasized *and* or used *both*	"only cat", "just elephant", "both!", "cat AND dog!"

Children's open-ended feedback. As explained in section , we also categorized and annotated children's spontaneous and free-form verbal reactions to the puppet's guesses. Table 5 summarizes the definitions and examples for each category and Figure 12 shows the results. We should point out that each trial type had a similar number of "None" cases. Some children remained more or less silent throughout the experiment and only provided rewards to the puppet. In the next study we ask children to provide feedback explicitly and therefore we have no "None" responses. In the discussion and analysis here we will not comment further on the "None" category but focus on the other three categories.

In the leftmost column, when the guessed animal was not on the card (e.g. elephant), children either provided judgments like "No!" or described what was on the card like *cat* or *cat and dog*. However, when the guessed animal was the only animal on the card (e.g. cat), most children provided a positive judgment like "Yes". When the animal guessed was only one of the animals on the card, children described what was on the card, for example, *cat and dog*. Corrections were rare for all these four trial types.

In the critical trial types with conjunction and disjunction, children showed a high rate of corrections and description when the guess used *and* but there was only one animal on the card. In their corrections, children used the focus particles *just* and *only* as in "just a cat" or "only a cat". However, in trial types where conjunction was used and both animals were depicted, children predominantly provided positive judgments like "Yes!" and "You are right". Considering disjunctive guesses like "cat or dog", when only one of the animals was

on the card, most children simply described what was on the card, for example "cat". However, when both animals were on the card, children corrected the puppet by saying "Both!" or emphasizing *and* as in "cat AND dog!"

We performed chi-squared goodness-of-fit tests to compare the feedback distributions in the critical conditions with *and* and *or*. Here we focus on those trials (the four bar charts on the right of Figure 12). Children's linguistic feedback showed three patterns. First, the one-animal conjunctive and two-animal disjunctive (top left and bottom right) trials contained a higher proportion of corrections than the other trial types. These were trials where the guesses were either false or infelicitous. In the conjunction trials, a comparison of the feedback distribution in one-animal and two-animal conditions was statistically significant ($\chi^2(3, 83) = 201.65$, $p < .0001$), suggesting that children gave different feedback to true and false guesses. A similar numerical trend was present in the disjunction trials, but it was not significant ($\chi^2(9, 4) = 12$, $p = 0.21$).

Second, the one-animal disjunctive trials (top right) showed the highest proportion of "descriptions". These are trials in which the guess is correct but not specific enough: it leaves two possibilities open. These trials were significantly different from the one-animal trials for conjunction ($\chi^2(3, 83) = 62.16$, $p < .0001$). Finally, the two-animal conjunctive trials (bottom left) showed the highest proportion of "judgments" such as *You are right!*. This was not surprising given that these trials represented the optimal guessing scenario. These trials had a significantly different feedback distribution from the matching disjunction trials ($\chi^2(3, 84) = 184.98$, $p < .0001$).

Discussion. In study 2, we used a 3AFC judgment task to test children's comprehension of logical connectives *and* and *or*. We compared these results to those found in the 3AFC judgment task of study 1 with adults. The general comparison showed that adults and children had similar patterns of judgments, except when both disjuncts were true. In such cases, adults judged the disjunctive guess as not completely right while most children found it completely right. There was even a slight preference among children to rewarded the

puppet more in such cases, compared to cases of disjunction when only one disjunct was true.

To consider another measure of children's comprehension, we also looked at children's spontaneous open-ended verbal feedback to the puppet's guesses. Our analyses suggested that children recognized false and infelicitous utterances with the connectives and provided appropriate corrective feedback. As expected from an adult-like understanding of connectives, children corrected the puppet most often when there was only one animal on the card and the guess was conjunctive, or when there were two animals on the card and the guess was disjunctive. Perhaps the most important finding was that children increased their corrective feedback in disjunctive guesses where both disjuncts were true, compared to those with only one true disjunct. These findings differ from the results of the 3AFC judgment task which suggested that children did not find any infelicity with disjunctive guesses when both disjuncts were true.

The analysis of children's open-ended feedback raises two important issues. First, as we mentioned before, it runs counter to what the 3AFC judgment task suggests with respect to exclusivity implicatures. The forced-choice task suggests that children find such underinformative utterances as unproblematic while analysis of their spontaneous feedback shows that they provided more corrections to such utterances. Second, a common explanation for why children fail to derive implicatures is that they cannot access the stronger alternative to the disjunction *or*, namely *and* (Barner, Brooks, & Bale, 2011). However, in the context of the guessing game, some children explicitly mentioned the word *and*, as the word the puppet should have said instead of *or*. Interestingly, these children continued to reward the puppet and considered the guess "right". This raises the possibility that children's forced-choice truth value judgments, whether with two or three alternatives, do not fully reflect their pragmatic knowledge. In study 3, we used both a 2AFC truth judgment task and an analysis of children's open-ended feedback. If the findings of study 2 were on the right track, we expected to replicate the same pattern in study 3, namely that the analysis of children's open-ended feedback should provide more evidence that children

are sensitive to pragmatic violations than the results of the 2AFC judgments.

Study 3: Children's 2AFC judgments and open-ended feedback

This study used the same paradigm as study 2 but focused on children's open-ended feedback and aimed at replicating the findings in study 2. The main hypothesis was that four-year-olds provide corrective feedback to the puppet if both disjuncts are true, but they do not consider this infelicity to be grave enough to render the guess itself "wrong". The main hypothesis along with relevant analyses and predictions were preregistered in an "As Predicted" format[5]. The study used a 2AFC judgment task to compare with the open-ended feedback results. The prediction was that children would provide corrective feedback to the puppet when both disjuncts were true, yet consider the guess "right" and not reflect this infelicity in their truth value judgments. This is what the study found.

Table 6

Summary of Study 1, 2, and 3 Methods

Study	N	Age	Mode	Response Options
Study 1 - Part 1	57	Adults	Online (Mturk)	Wrong, Right
Study 1 - Part 2	52	Adults	Online (Mturk)	Wrong, Kinda Right, Right
Study 2	42	3;1-5;2 (M = 4;3)	Study Room	Circle (Wrong), Little Star (Little Right), Big Star (Right)
Study 3	50	3;6-5;9 (M = 4;7)	Study Room	Yes (Right)/No (Wrong) - Open-ended Feedback

Methods.

[5]The As Predicted pdf document is accessible at https://aspredicted.org/x9ez2.pdf.

Materials and Design. Study 3 was similar to Study 2 but differed in how children provided their judgments. Based on the findings in Study 2, we focused on verbal judgments and feedback, instead of rewards. We used two different ways of measuring children's judgments. First, we encouraged children to provide verbal feedback to the puppet. They were asked to say "yes" when the puppet was right, and "no" when he was not. They were also encouraged to help the puppet say it better when he was not right. After children were done with this initial open-ended feedback, for each trial we asked a forced choice yes/no judgment question: "Was Jazzy (the puppet) right?". This question elicited a "yes" or "no" response for each trial independent of their earlier open-ended response. These two measures allowed me to compare open-ended and forced-choice judgments.

Participants and Procedure. We recruited 50 English speaking children from the Bing Nursery School at Stanford University. Children were between 3;6 and 5;9 years old (Mean = 4;7). The setup and procedure were similar to Study 2, except there were no rewards on the table. As before, participants sat through three phases: introduction, instruction, and test. The introduction phase made sure children knew the names of the animals on the cards. In the instruction phase, they received four training trials, as shown in Table 7.

As in Study 2, the experimenter put a sleeping mask over the puppet's eyes and explained that Jazzy (the puppet) was going to guess what animal was on the cards. He then picked the first card and asked the puppet: "*What do you think is on this card?*" The puppet replied with "*There is a dog*". The experimenter showed the cat-card to the child and said: when Jazzy is *not right*, tell him "no". He then asked the child to say "no" to the puppet. The second trial followed the same pattern except that the puppet guessed *right* and the experimenter invited the child to say "yes" to the puppet. There were two more instruction trials before the test phase began. This contained 16 randomized trials, half of which contained guesses with the words *and* and *or*. The randomization code as well as the details of the methods are on the online repository for this dissertation at

https://github.com/jasbi/jasbi_dissertation_LearningDisjunction.

Table 7

Instruction Trials for Study 3.

Card	Guess	Response
CAT	there is a dog!	No!
ELEPHANT	there is an elephant!	Yes!
DOG-ELEPHANT	there is a cat!	No!
DOG	there is a dog!	Yes!

Results. We first look at the results of the 2AFC judgement task for each trial type and compare them to those of the adults' in Study 1. Then we analyze children's open-ended responses and compare them to the forced choice responses obtained in the same trial types. For the 2AFC judgments we excluded 26 trials (out of total 800) where children either did not provide a Yes/No response or provided both (i.e. "Yes and No"). The exclusions were almost equally distributed among different types of guesses and cards. In the analysis of children's open-ended feedback, we excluded 8 trials (out of total 800) where children either did not provide any feedback or their feedback could not be categorized into the existing categories.

Two-Alternative Forced Choice Judgments. Figure 13 shows children's 2AFC judgments. In the leftmost column, when the animal guessed was not on the card (e.g. elephant), children considered the guess "wrong". When the animal guessed was the only animal on the card (e.g. cat), children considered the guess "right". However, if the animal guessed (e.g. cat) was only one of the animals on the card, children were equally split between "wrong" and "right" judgments. On the other hand, almost all adults considered such guesses "right" in their 2AFC judgments (Figure 4). In such trial types, children seem to interpret the guess "there is a cat" as "there is **only** a cat", while adults do not. This difference between children and adults is unexpected for a theory of meaning acquisition that

assumes children are overall more logical or literal as interpreters than adults (Noveck, 2001).

In the trials with *and* and *or*, children's judgments were similar to those of adults. Figure 14 compares adults' and children's 2AFC judgments. In trials with conjunction, when only one of the animals was on the card, most children considered the guess "wrong". This is similar to adults' judgments, but different in extent: adults were more consistent and unanimous in rejecting such guesses. A mixed effects logistic regression with the fixed effect of age category (adult vs. child) and random effect of subject found no significant difference between adults' and children's responses in such trials (see Table 8, Conjunction - One Animal).

Table 8

Mixed effects logistic models for conjunction and disjunction trials when only one disjunct was true, in 2AFC judgments of adults and children, using `glmer` *in R's lme4 package. Formula:* $Response \sim AgeCategory + (1|Subject)$.

Trial Data	Coefficient	Standard Error	Z-Value	P-value
Conjunction - One Animal	-2.05	2.86	-0.72	0.47
Disjunction - One Animal	1.34	1.79	0.75	0.45

In conjunctive guesses where both animals were on the card, both children and adults were unanimous in considering the guess "right". In disjunctive trials when only one of the animals was on the card, most children considered the guess "right". This is again similar to adults but differs from them in extent: adults more consistently and unanimously judged such guesses as "right". Yet again, a mixed effects logistic regression with the fixed effect of age (adult vs. child) and random effect of subject found no significant difference between adults' and children's responses in such trials (see Table 8, Disjunction - One Animal). Adults and children showed almost identical patterns of judgments in trials where there was two animals on the card and the guess used the connective *or*. Children and adults did not differ in their rate of rejecting disjunctive guesses when both disjuncts were true.

Finally, there is a small but significant preference in children's judgments of disjunctive statements for both disjuncts to be true. Comparing the disjunctive trials with one animal and two animals on the card, a mixed-effects logistic model with the fixed effect of disjunction type and the random effect of subjects found that children had a slight preference for both animals to be on the card (b= 1.85, se= 0.56, z= 3.32, $p < 0.001$). There was a similar small trend in children's three-alternative judgments in study 2. While this was quite small compared to the other effects observed in these studies, it nevertheless indicated a difference between children's and adults' judgments. We return to this in more detail in section of the General Discussion.

Open-ended Feedback. Figure 15 shows the distribution of children's feedback to the puppet in Study 3 (see Table 5 for the definitions and examples of feedback categories). There were no "None" responses in this study since the experimenter explicitly asked children to provide feedback to the puppet. The distribution of the responses in the other three categories (Judgment, Description, and Correction) revealed a successful replication of Study 2.

Children's feedback showed four main patterns. First when the puppet guessed an animal not on the card (e.g. *There is an elephant!*), there is a split pattern between negative judgments like *No!* and simply mentioning the animal on the card, e.g. *Cat!*. Children provided no corrections on such trials, at leat the way we have defined them. Second, almost all children responded with positive judgments like *Yes!* when the puppet's guess accurately matched what was on the card. This was the case in trials where there was only one animal on the card (e.g. cat) and the puppet mentioned it (e.g. *There is a cat!*), as well as trials where there were two animals on the card and the puppet mentioned both with a conjunction (e.g. *There is a cat and a dog!*). Third, children provided the largest number of corrective feedback in trials where the guess was either false or infelicitous. These included three trial types: (a) the ones where there were two animals on the card (e.g. cat and dog) but the puppet only guessed one (e.g. *There is a cat!*); (b) the ones where the puppet

guessed two animals with conjunction (e.g. *There is a cat and a dog!*) but only one of them was on the card (e.g. cat); and (c) the ones where there were two animals on the card (e.g. cat and dog), and the puppet guessed both but used a disjunction (e.g. *There is a cat or a dog!*). Finally, there was a pattern of feedback unique to disjunctive trials (e.g. *There is a cat or a dog!*) with only one animal on the card (e.g. cat). In such cases, almost all children simply named the animal on the card (e.g. *Cat!*).

Figure 16 breaks down children's open-ended feedback based on whether children said *Yes!*, *No!*, or said something else. Responses that were not yes/no judgments are grouped in a middle category shown with a dash. The goal here is to compare children's open-ended judgments with their forced choice judgments shown in Figure 13. Children's open-ended judgments and their forced choice judgments in study 3 show similar patterns for all types of guesses except for disjunctive ones. In trials that the puppet guessed with *or*, the vast majority of children refused to provide a yes/no judgment when they were not forced to. Instead, they described the animal on the card or provided corrections to the puppet's infelicitous disjunctive guess.

One way to interpret these results is that disjunctive guesses (with at least one disjunct true) are considered neither right nor wrong by almost all children. When children were forced to provide wrong/right responses in the experimental context, some conformed to the adult patterns of judgment and some did not. However, it is possible that such deviations from adult judgments do not reflect differences in the comprehension of disjunction, but rather differences in how children map their comprehension of disjunction onto the notions of "right" and "wrong" in a forced choice judgment task. In other words, it is possible that children and adults only differ in how they behave when they are forced to respond with a fixed set of options.

Figure 17 shows the proportion of feedback categories other than yes/no judgments on the x-axis. My goal here is to display the trial types with corrective feedback (blue and red). These trial types include: (1) conjunction when only one conjunct is true (e.g. guess: *There*

is a cat and a dog!, card: cat), (2) disjunction when both disjuncts are true (e.g. guess: *There is a cat or a dog*, card: cat and dog), and (3) simple guesses when two animals were on the card (e.g. *There is a cat!*, card: cat and dog). These trial types involved guesses that were either false or infelicitous. Furthermore, the type of corrective feedback children provided matched the type of mistakes made in the guesses. With conjunctive guesses (e.g. There is a cat and a dog¡‘) when there was only one animal on the card (e.g. cat), children provided exclusive corrections (e.g. *Just/only a cat!*), suggesting that the other animal (e.g. dog) should have been excluded. When two animals were on the card (e.g. cat and dog) and the puppet used a disjunctive guess (e.g. *There is a cat or a dog!*), or simple guess (e.g. *Ther is a cat!*), children provided inclusive feedback, suggesting that another animal should have been included. This is particularly notable in the case of disjunction since both animals were mentioned, but children still emphasized that the connective *and* should have been used, or that both animals mentioned were on the card.

Discussion. Study 3 measured children's comprehension of logical connectives in two ways: First, with analyzing their open-ended feedback and second, with a two-alternative forced choice task. First, we asked children to say *yes* to the puppet if he was right and *no* if he was wrong. However, children could provide any form of feedback they wanted. Second, we followed children's open-ended feedback with a two-alternative forced choice question: *Was the puppet right?* This way, we could measure children's comprehension in two different ways in the same trial. Ideally, both measures should show similar results. However, the findings were similar for conjunctive guesses, but not disjunctive ones. Children avoided binary right/wrong feedback with disjunction and preferred to provide more nuanced feedback.

The 2AFC responses followed the predicted pattern: conjunctive guesses were judged wrong if only one conjunct was true, and right if both were true. Disjunctive guesses were judged right whether one or both disjuncts were true. There was no significant difference in the 2AFC task between the responses of children and those of adults in Study 1.

Children's open-ended feedback in Study 3 replicated the findings of Study 2. Children provided more corrective feedback in false and infelicitous trials than in true and felicitous ones. The corrective feedback was tailored to the puppet's mistake. If the puppet used a conjunction when there was only one animal on the card, children pointed out that the other animal should have been excluded from the guess. They used the exclusive adverbials *just* and *only* in their feedback. If the puppet used a disjunction when both animals were on the card, children stressed *and* or *both*, implying that both animals should have been included.

While the 2AFC results suggested that children took no issue with disjunctive guesses when both disjuncts are true, the analysis of their corrective feedback showed that they provide appropriate corrections in such cases and emphasize that the connective *and* would have been a better guess. Taking both measures together, we conclude that even though children are aware of the problem with such guesses, they do not consider them *wrong*. These results are similar to those we reported for adults in Study 1.

General Discussion

We reported three studies on adults and four-year-olds' comprehension of the logical connectives *and* and *or*. The first study used two- and three-alternative forced choice judgment tasks with adults. In the 2AFC task, adult interpretations closely matched the semantic accounts of *and* and *or* as conjunction and inclusive disjunction. The 2AFC judgments did not register robust signs of pragmatic infelicities. However, the 3AFC judgments showed signs of pragmatic infelicities, especially in disjunctive guesses with true disjuncts. When two animals where on the card (e.g. cat and dog) and the guess used *or* (e.g. *There is a cat or a dog!*), participants were more likely to choose "kinda right" rather than "right".

The second study used a 3AFC judgment task with four-year-old children. It also included an exploratory analysis of children's open-ended verbal feedback to the puppet in the experimental setting. Children's interpretations were similar to those of adults in the

3AFC task and only differed for pragmatically infelicitous disjunctions. When both disjuncts were true, adults tended to judge disjunctive guesses as "kinda right". This was evidence for the pragmatic infelicity of such guesses. While, children judged such disjunctive statement as "right", the analysis of their open-ended feedback showed that they took issue with such statements as well, and provided appropriate corrective feedback.

In the third study, we focused on eliciting open-ended verbal feedback from children and followed it with a 2AFC task. In the 2AFC task, children's responses reflected the semantics of connectives as conjunction and inclusive disjunction. There was no significant difference between children and adults in the two-alternative judgments. Since the 2AFC task appeared to be a good indicator of semantic knowledge, it seemed reasonable to conclude that adults and four-year-olds displayed similar semantic knowledge of the connectives. Analysis of the children's open-ended feedback replicated the findings in study 2. Children provided more corrective feedback in false and pragmatically infelicitous trials with logical connectives than in felicitous trials. The comparison of the 2AFC task and children's open-ended responses showed that children are sensitive to the infelicity of disjunctions with true disjuncts, even though they consider them to be "right" guesses.

Overall, we did not find any major differences between adults' and four-year-old children's interpretations of logical connectives *and* and *or* in the context of the guessing game. However, there were two minor differences. First, we found that in both 2AFC and 3AFC judgment tasks, children showed a small preference for disjunctions with both disjuncts true rather than only one. Adults on the other hand showed the opposite pattern: they preferred disjuncts with only one disjunct true. Second, in both 2AFC and 3AFC judgment tasks, children rated disjunctions with both disjuncts true higher than adults did. That is, they considered utterances like *There is a cat or a dog* when both animals were on the card "right" more often than adults did. Here we will discuss these two differences and their possible causes in more detail.

Preference for True Disjuncts. First for some children, there was a small preference for both disjuncts being true, compared to only one. This effect is similar in kind but not magnitude, to an effect that Singh et al. (2016) and Tieu et al. (2016) reported. In our study this effect is quite small while Singh et al. (2016) and Tieu et al. (2016) seem to have found bigger effects. Based on this, Singh et al. (2016) proposed that many children at this age-range have a pragmatically driven conjunctive interpretation of disjunction. In short, due to a non-adult like alternative set to the connective *or*, children strengthen a disjunctive statement pragmatically and derive a conjunction. The studies reported here provide no support for this proposal. In both 2AFC and 3AFC judgments, children clearly differentiated between disjunctive and conjunctive guesses. Furthermore, analysis of children's open-ended feedback showed distinctly different response patterns for conjunction and disjunction. More importantly, the open-ended feedback to disjunctive guesses showed the opposite pattern to that predicted by the conjunctive hypothesis. Children took issue with disjunctions that had both disjuncts true and provided more corrective feedback in such cases. Therefore, the findings from Singh et al. (2016) and Tieu et al. (2016) may be a product of experimental design rather than a real reflection of children's comprehension of the connectives.

However, even if this small preference for true disjuncts is not due to the method of measurement, it can be accounted for in several other ways that have not yet been successfully ruled out. First, the conjunctive interpretation may not be due to a faulty pragmatic computation, but rather a default conjunctive interpretation when the connective is not properly heard, understood or is unknown. To check this hypothesis, it should be possible to test children's comprehension of novel or noisy connectives. A novel coordination like *cat dax dog* with *dax* as a nonce connective could well be interpreted as a conjunction. Such a result would suggest that in studies with high cognitive demand, children may default to a conjunctive interpretation if they miss the relevant connective. Second, the conjunctive preference could be due to some children's preference for the linguistic labels to match the animals on the card (or more generally a match between linguistic description and the state

of the world). This hypothesis is consistent with the results in the other trial type that had a mismatch in the number of animals and the guess, where the guess was still technically true: simple guesses (e.g. there is a cat) with two animals (e.g. cat and dog). Children were equally split between "wrong" and "right" in their judgments here, while adults considered such guesses "right". In light of these alternative explanations, we are hesitant to attribute this small preference to a pragmatically driven conjunctive interpretation of disjunction.

Lack of infelicity with true disjuncts in the forced choice tasks. The second difference between adults and children emerged in the 3AFC judgment task: in disjunctive trials (e.g. *There is a cat or a dog*) with two animals (e.g. cat and dog), adults were more likely to choose "kinda right" than children were. Children mostly chose "right". This response pattern has been taken to mean that children found no infelicity with such disjunctions or that they did not "derive an exclusivity implicature". The absence of an infelicity/implicature is consistent with the generalization that children are more likely than adults to interpret scalar terms literally, and that children do not compute implicatures or judge infelicity to the same **rate** that adults do (Pouscoulous & Noveck, 2009, Katsos (2014)). But why is that?

There have been at least three major proposals to account for children's low rate of implicatures: 1. processing difficulty (Pouscoulous, Noveck, Politzer, & Bastide, 2007; Reinhart, 2004) 2. non-adult-like lexical entry (Barner et al., 2011; Horowitz, Schneider, & Frank, 2017) and 3. pragmatic tolerance (Katsos & Bishop, 2011). Here we show that none of these accounts can provide a satisfactory explanation of the results in this study.

1. Processing difficulty. First, processing accounts locate the problem in children's processing capacities such as working memory. They suggest that pragmatic computations are cognitively taxing and children lack the appropriate processing resources to carry them out appropriately. A prediction of processing accounts (at least in their current format) is that children will show reduced implicature computations for all types of implicatures – scalar or ad-hoc. This prediction was not borne out in our experimental results here. In

Study 3, children were much more likely than adults to call a simple guess (e.g. *There is a cat!*) "wrong" if there were two animals on the card (e.g. cat and dog). Processing accounts do not predict that children may derive implicatures at a higher rate than adults but this is what we found, at least for the traditional interpretation of the judgment task.

2. Non-adult-like Lexicon. Several proposals blame the structure of the child's lexicon for the alleged failure in deriving implicatures. The assumption is that the child's lexical entry for scalar items must include three elements for successful derivation: 1. the semantics of the weak term (e.g. *some*, *or*) 2. the semantics of the strong term (e.g. *all*,*and*); and possibly 3. a scale that recognizes the stronger term as an alternative to the weaker one (e.g. <*some*, *all*>, <*or*, *and*>). Each of these elements have been pinpointed as the source of the problem in previous studies (Barner et al., 2011; Horowitz et al., 2017; Katsos & Bishop, 2011). However none of them seem to apply to the results reported here.

If children in this study lack the semantics of the connective *or*, we would expect them to either perform at chance or default to a conjunctive interpretation. Neither prediction was borne out in studies 2 and 3. Furthermore, children's free-form linguistic feedback in both studies suggested that children understood disjunction well enough to provide relevant feedback. So this explanation seems unlikely. The problem cannot be that children do not know the meaning of *and* either. Children's performance in both study 2 and 3 for conjunction trials show that they understand its meaning very well. Finally, while it is possible that children lacked the appropriate lexical scale and could not access the stronger alternative, this explanation cannot be the whole story. Several children in both studies stressed the word *and* in their verbal feedback, suggesting that the puppet should have used the stronger term instead. However, they still judged the puppet's guess as "right". If children could not access the stronger term, they could not mentioned it in their feedback either.

3. Pragmatic Tolerance. Katsos & Bishop (2011) suggested that children tend to tolerate pragmatic infelicities more than adults. They showed that when children were

provided with a 2AFC judgment task, they considered a description with the scalar term *some* as "right" when *all* was more informative (e.g. *The turtle played with some of the balls.*, Scene: the turtle played with all the balls.) However, when they are presented with three options (small, big, and huge strawberries) in a 3AFC task, they choose the middle option in the same type of trials. They argue that children tolerate pragmatic infelicities and do not regard them as "wrong". As in a processing account, the tolerance account predicts that scalar and ad-hoc implicatures will be similarly affected. However, our results did not match those of Katsos & Bishop (2011). When children were presented with a 3AFC task, they chose the highest reward (and not the middle option) for uses of *or* when *and* was more informative. Second, and more importantly, we found different patterns for exhaustive and scalar inferences as mentioned before. This is not predicted by the tolerance account unless we assume that children are more tolerant towards violations of scalar inferences than they are towards exhaustive ones. While this is not currently assumed in the literature, it is a possible adjustment. However, we would like address this issue by focusing on another related factor: the role of measurement in estimates of children's pragmatic capacity (Katsos, 2014). Several observations in the current studies provide support for the hypothesis that methodological issues, and more specifically issues of measurement contribute to the differences found between adults and children in pragmatic capacity. First, Study 1 showed that even for adults, the estimates of adult infelicity rates may differ based on the number of alternatives in the forced choice task. A 2AFC task underestimated adults' sensitivity to pragmatic infelicity. In fact, in a follow up study, we systematically varied the number of response options and replicated the results presented here (see Jasbi, Waldon, and Degen in press). Second, children's open-ended linguistic feedback in the experimental context better reflected their sensitivity to pragmatic nuances than the forced-choice judgment tasks. Third, children showed a higher rate of infelicity judgments for cases of ad-hoc implicatures (simple guesses with two animals on the card) than adults did. While a difference in sensitivity to ad-hoc vs. scalar implicatures has been reported and argued for before (Horowitz et al., 2017;

Stiller, Goodman, & Frank, 2015), a higher sensitivity than adults is not predicted by any of the current accounts.

In order to better understand the differences between adults and children's pragmatic capacities, it is necessary to have a good understanding of how our measurements affect estimates of adults and children's performance in the experimental tasks. Children may be no more capable of making exhaustive inferences than adults and no less capable of making scalar inferences either. They may simply have a different construal of the wrong-right scale and of what the forced-choice task is about. The concepts "right" and "wrong" are as much subject to developmental change and differences between adults and children as are scalar items that constitute the focus of our studies. It is possible that children's understanding of what constitutes as "right" or "wrong" does not fully conform to that of adults. However, it remains to be established what these differences are and how they affect the estimates of children's pragmatic abilities. It is important to point out that such issues of measurement could be the culprit behind both children's seemingly slight preference for true disjuncts described earlier and the lack of infelicity judgments when both disjuncts are true.

A General Approach for Measuring Implicature/Infelicity Rate

Methodological issues are nothing new in developmental studies and language acquisition. Creating better measures of children's linguistic capacities has always been a major concern for researchers in the field. Our goal here is to propose some future steps that can address possible methodological issues in assesssing children's pragamtic competence.

As Pouscoulous & Noveck (2009) and Katsos (2014) have suggested, the central issue is "the rate" at which children and adults manifest pragmatic reasoning in the experimental setting. No one doubts children's capacity to perform such computations. At issue is the extent to which children and adults compute specific implicatures. The claim is that children perform such computations less often than adults; or that children do not perform such computations where adults normally do. In the previous section, we discussed some factors that might account for these differences including processing demands, the structure of the

lexicon, tolerance, as well as issues of measuring adults and children's comprehension. As Katsos (2014) pointed out, it seems reasonable to assume that all these factors play some part here. What matters is the degree to which each contributes to the outcome.

Figure 18 shows the factors that affect pragmatic computations as well as the observations of the rate of pragmatic computations in an experiment. First it is important to distinguish between factors that affect pragmatic computations and those that affect the observed rate in an experimental setting. As we showed in Study 1, given the number of alternatives in the forced choice task (2AFC vs. 3AFC), we may get different estimates of adults' rate of infelicity judgments, but we cannot to assume that there is a difference in adults' pragmatic capacities in these two tasks. A similar situation exists when we compare children's forced choice measures of infelicity and their open-ended feedback. In disjunctive trials where both disjuncts are true, the forced choice tasks show no sign of children detecting infelicity while the open ended responses show that children are sensitive to the infelicity of disjunction when a conjunction would have been more appropriate.

Conclusion

To conclude, the studies presented here did not provide evidence for a substantial difference between adults and three-to-five-year-old children in their **semantic** knowledge of the logical connectives *and* and *or*. The results were highly consistent with the current accounts that posit the semantics of *and* as conjunction and *or* as inclusive disjunction. With respect to pragmatic knowledge, the three-alternative forced choice judgment task showed that adults are sensitive to the infelicity of disjunctive statements when both disjuncts are true. We also showed that the three-alternative judgment task failed to register such a sensitivity for children, but our systematic analysis of children's open-ended verbal feedback did. It showed that children can provide appropriate corrections to infelicitous utterances containing logical connectives *and* and *or*.

References

Aloni, M. (2016). Disjunction. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Stanford University. Retrieved from https://plato.stanford.edu/archives/win2016/entries/disjunction/

Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition*, *118*(1), 84–93.

Braine, M. D., & Rumain, B. (1981). Development of comprehension of "or": Evidence for a sequence of competencies. *Journal of Experimental Child Psychology*, *31*(1), 46–70.

Chierchia, G., Crain, S., Guasti, M. T., & Thornton, R. (1998). "Some" and "or": A study on the emergence of logical form. In *Proceedings of the Boston University conference on language development* (Vol. 22, pp. 97–108). Somerville, MA: Cascadilla Press.

Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A., & Meroni, L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In *Proceedings of the 25th Boston University conference on language development* (pp. 157–168). Somerville, MA: Cascadilla Press.

Chierchia, G., Guasti, M. T., Gualmini, A., Meroni, L., Crain, S., & Foppolo, F. (2004). Semantic and pragmatic competence in children's and adults' comprehension of or. In I. Noveck & D. Sperber (Eds.), *Experimental pragmatics* (pp. 283–300). Basingstoke: Palgrave Macmillan.

Clark, E. V. (1973). Non-linguistic strategies and the acquisition of word meanings. *Cognition*, *2*(2), 161–182.

Crain, S. (2012). *The emergence of meaning*. Cambridge: Cambridge University Press.

Crain, S., & Khlentzos, D. (2008). Is logic innate? *Biolinguistics*, *2*(1), 024–056.

Crain, S., & Khlentzos, D. (2010). The logic instinct. *Mind & Language*, *25*(1), 30–65.

Crain, S., & Thornton, R. (1998). *Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics*. Cambridge, MA: MIT Press.

Crain, S., Gualmini, A., & Meroni, L. (2000). The acquisition of logical words. *LOGOS and*

Language, *1*, 49–59.

Gazdar, G. (1979). *Pragmatics: Implicature, presupposition, and logical form*. New York: Academic Press.

Goro, T., & Akiba, S. (2004). The acquisition of disjunction and positive polarity in Japanese. In *Proceedings of the 23rd West Coast conference on formal linguistics* (pp. 251–264). Somerville, MA: Cascadilla Press.

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

Gualmini, A., & Crain, S. (2002). Why no child or adult must learn de Morgan's laws. In *Proceedings of the Boston University conference on language development*. Somerville, MA: Cascadilla Press.

Gualmini, A., Crain, S., & Meroni, L. (2000). Acqisition of disjunction in conditional sentences. In *Proceedings of the boston university conference on language development*.

Gutzmann, D. (2014). Semantics vs. pragmatics. In L. Matthewson, C. Meier, H. Rullmann, & T. E. Zimmermann (Eds.), *The companion to semantics*. Oxford: Wiley.

Horn, L. (1989). *A natural history of negation*. Chicago, IL: University of Chicago Press.

Horowitz, A. C., Schneider, R. M., & Frank, M. C. (2017). The trouble with quantifiers: Exploring children's deficits in scalar implicature. *Child Development*.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures* (Vol. 84). London: Routledge.

Jasbi, M., Waldon, B., & Degen, J. (submitted). *Linking hypothesis and number of response options modulate inferred scalar implicature rate*.

Johansson, B. S., & Sjolin, B. (1975). Preschool children's understanding of the coordinators "and" and "or". *Journal of Experimental Child Psychology*, *19*(2), 233–240.

Katsos, N. (2014). Scalar implicature. In D. Matthews (Ed.), *Pragmatic development in first language acquisition* (Vol. 10, p. 183—198). Amsterdam: John Benjamins.

Katsos, N., & Bishop, D. V. (2011). Pragmatic tolerance: Implications for the acquisition of

1169 informativeness and implicature. *Cognition*, *120*(1), 67–81.

1170 Neimark, E. D. (1970). Development of comprehension of logical connectives: Understanding
1171 of "or". *Psychonomic Science*, *21*(4), 217–219.

1172 Neimark, E. D., & Slotnick, N. S. (1970). Development of the understanding of logical
1173 connectives. *Journal of Educational Psychology*, *61*(6p1), 451.

1174 Nitta, N., & Nagano, S. (1966). Basic logical operations and their verbal expressions: Child's
1175 conception of logical sum and product. *Research Bulletin of the National Institute for*
1176 *Educational Research, Tokyo*, *7*, 1–27.

1177 Notley, A., Thornton, R., & Crain, S. (2012). English-speaking children's interpretation of
1178 disjunction in the scope of "not every". *Biolinguistics*, *6*(1), 32–69.

1179 Notley, A., Zhou, P., Jensen, B., & Crain, S. (2012). Children's interpretation of disjunction
1180 in the scope of "before": A comparison of English and Mandarin. *Journal of Child*
1181 *Language*, *39*(03), 482–522.

1182 Noveck, I. A. (2001). When children are more logical than adults: Experimental
1183 investigations of scalar implicature. *Cognition*, *78*(2), 165–188.

1184 Paris, S. G. (1973). Comprehension of language connectives and propositional logical
1185 relationships. *Journal of Experimental Child Psychology*, *16*(2), 278–291.

1186 Pouscoulous, N., & Noveck, I. A. (2009). Going beyond semantics: The development of
1187 pragmatic enrichment. In S. Foster-Cohen (Ed.), *Language acquisition* (pp. 196–215).
1188 Berlin: Springer.

1189 Pouscoulous, N., Noveck, I. A., Politzer, G., & Bastide, A. (2007). A developmental
1190 investigation of processing costs in implicature production. *Language Acquisition*,
1191 *14*(4), 347–375.

1192 Reinhart, T. (2004). The processing cost of reference set computation: Acquisition of stress
1193 shift and focus. *Language Acquisition*, *12*(2), 109–155.

1194 Singh, R., Wexler, K., Astle-Rahim, A., Kamawar, D., & Fox, D. (2016). Children interpret
1195 disjunction as conjunction: Consequences for theories of implicature and child

development. *Natural Language Semantics*, *24*(4), 305–352.

Skordos, D., Feiman, R., Bale, A., & Barner, D. (2018, July). *Do children interpret "or" conjunctively?* Retrieved from https://osf.io/2srxk/

Stiller, A. J., Goodman, N. D., & Frank, M. C. (2015). Ad-hoc implicature in preschool children. *Language Learning and Development*, *11*(2), 176–190.

Su, Y. (2014). The acquisition of logical connectives in child Mandarin. *Language Acquisition*, *21*(2), 119–155.

Su, Y., & Crain, S. (2013). Disjunction and universal quantification in child mandarin. *Language and Linguistics*, *14*(3), 599–631.

Suppes, P., & Feldman, S. (1969). *Young children's comprehension of logical connectives. ERIC*. Department of Health, Education, Welfare. Office of Education.

Tarski, A. (1941). *Introduction to logic and to the methodology of the deductive sciences*. Oxford University Press.

Tieu, L., Yatsushiro, K., Cremers, A., Romoli, J., Sauerland, U., & Chemla, E. (2016). On the role of alternatives in the acquisition of simple and complex disjunctions in french and japanese. *Journal of Semantics*.

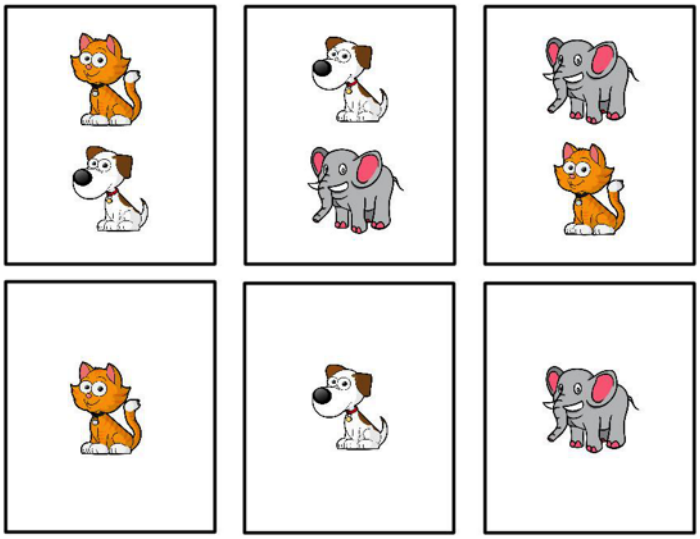

Figure 1. Cards used in the connective guessing game.

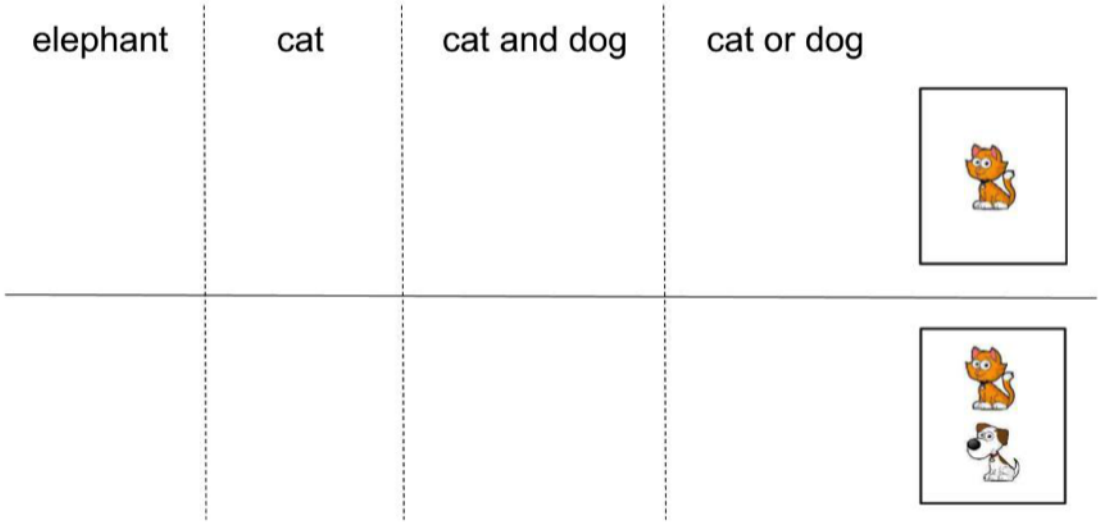

Figure 2. Trial types represented by example cards and example guesses.

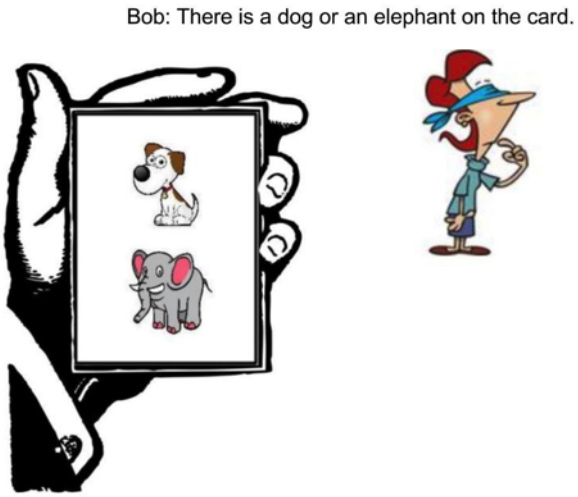

Figure 3. An example trial in Study 1.

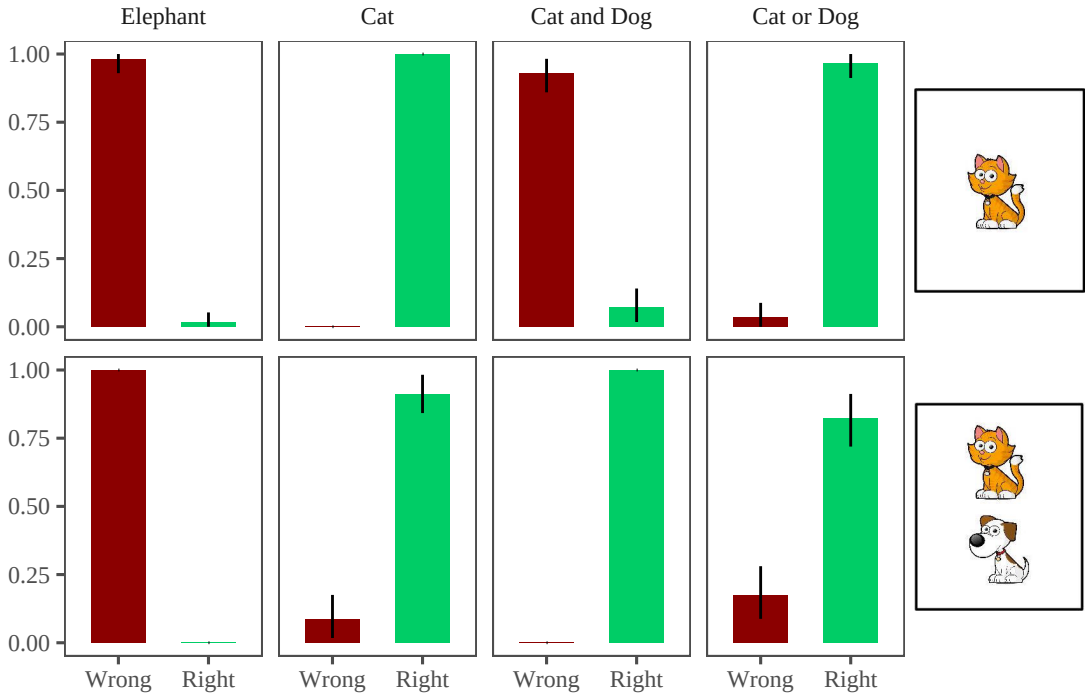

Figure 4. Adults' two-alternative forced choice judgments.

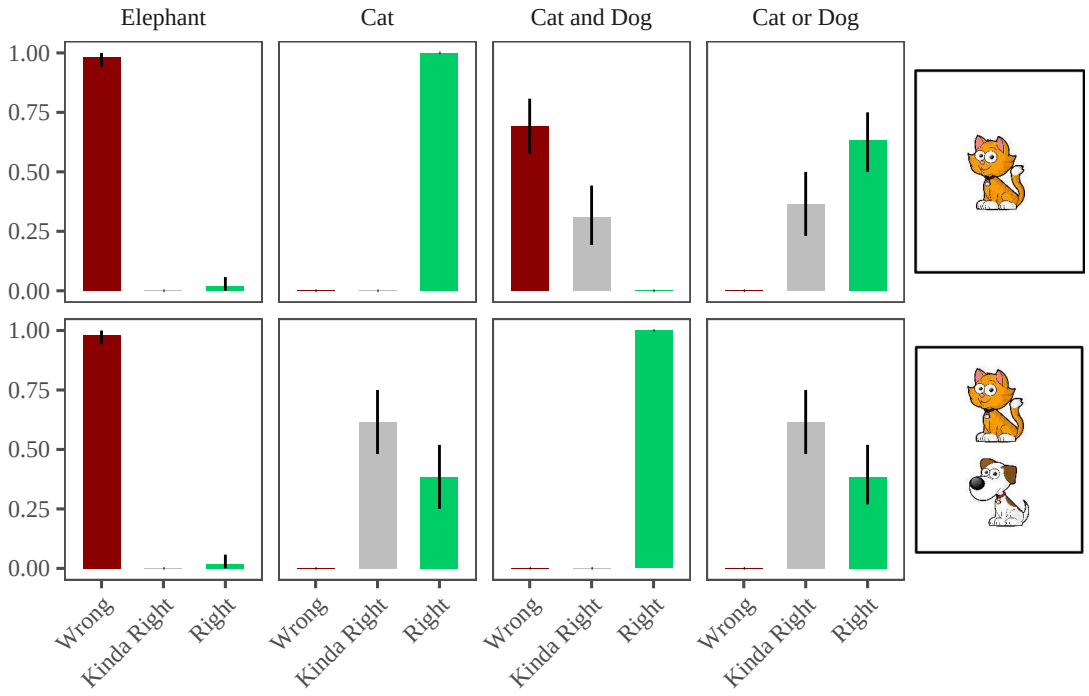

Figure 5. Adults' three-alternative forced choice judgments in the connective guessing game.

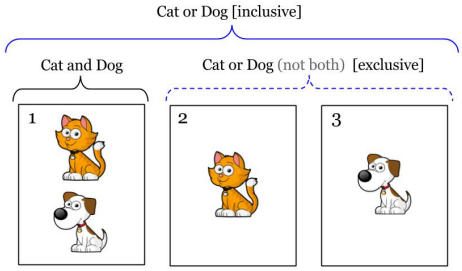

Figure 6. Example of cards referred to by a conjunction, inclusive disjunction, and exclusive disjunction.

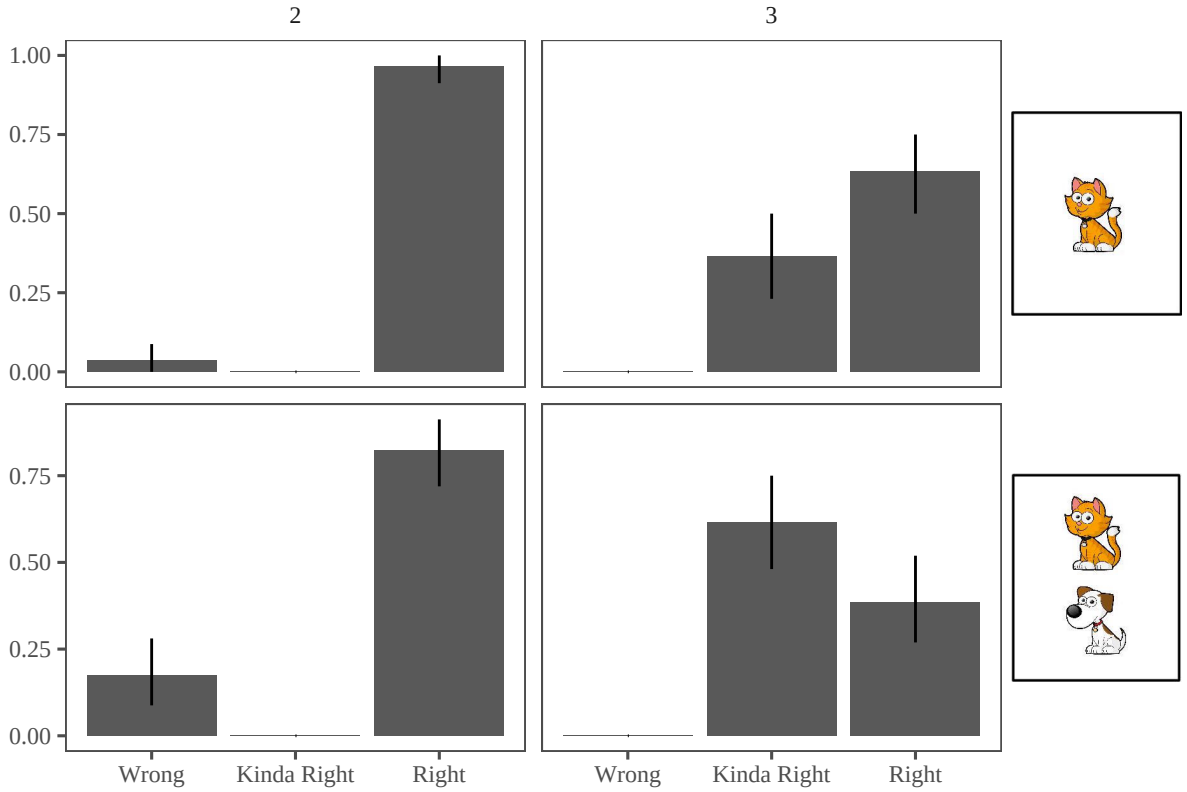

Figure 7. Adult responses to disjunction guesses like *cat or dog* with 2 and 3 options.

Figure 8. The puppet, Jazzy, with and without the sleeping mask.

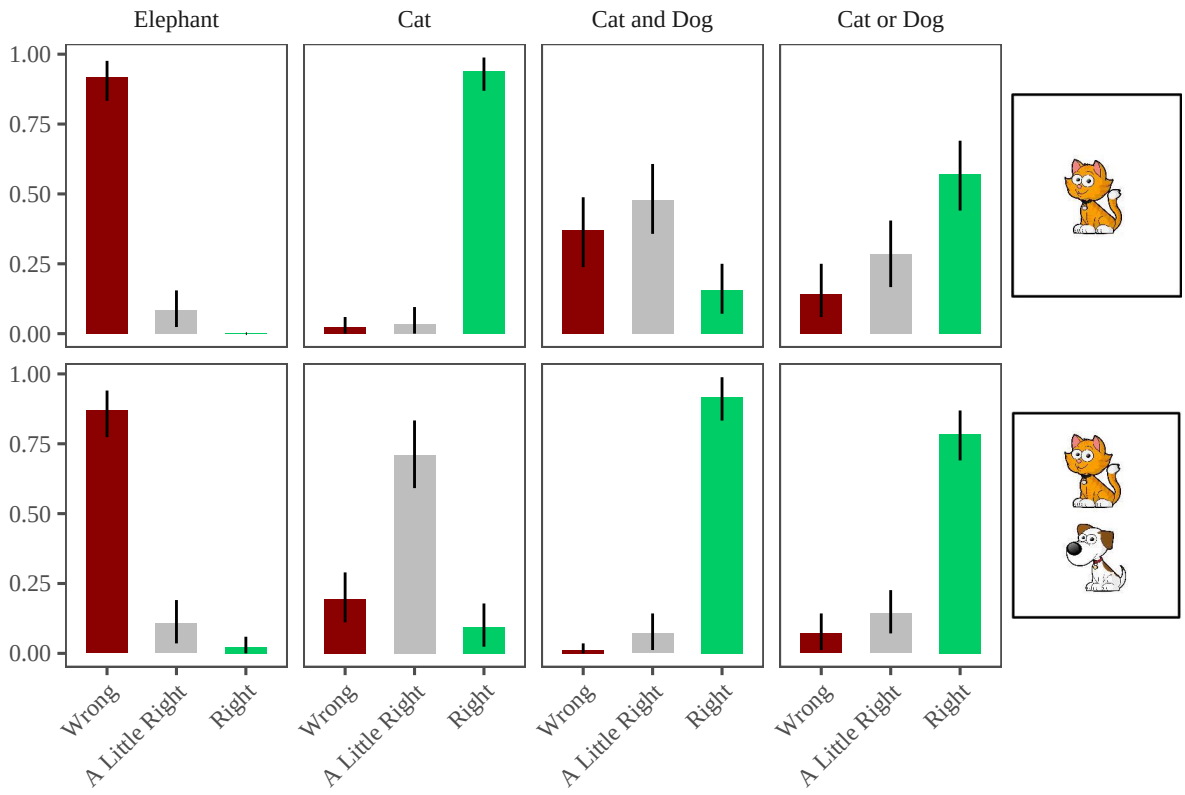

Figure 9. Children's 3AFC judgments in the connective guessing game.

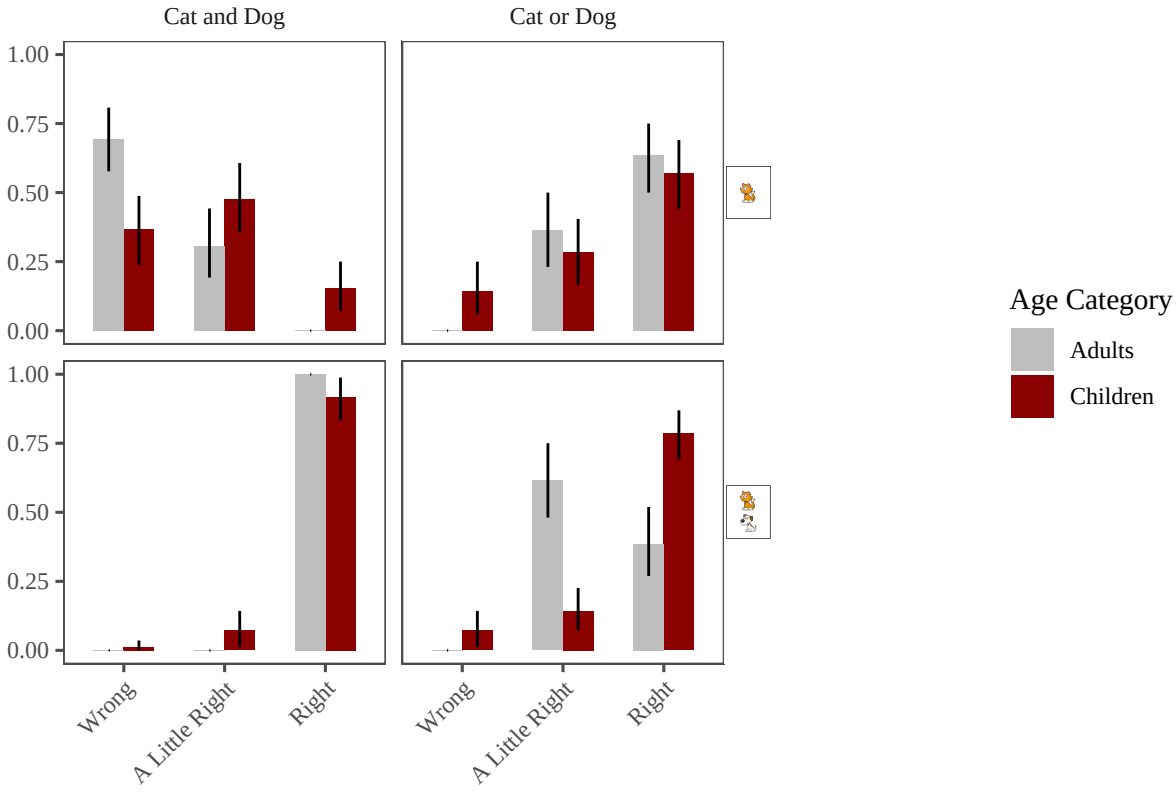

Figure 10. Comparison of Adults' and Children's 3AFC judgments.

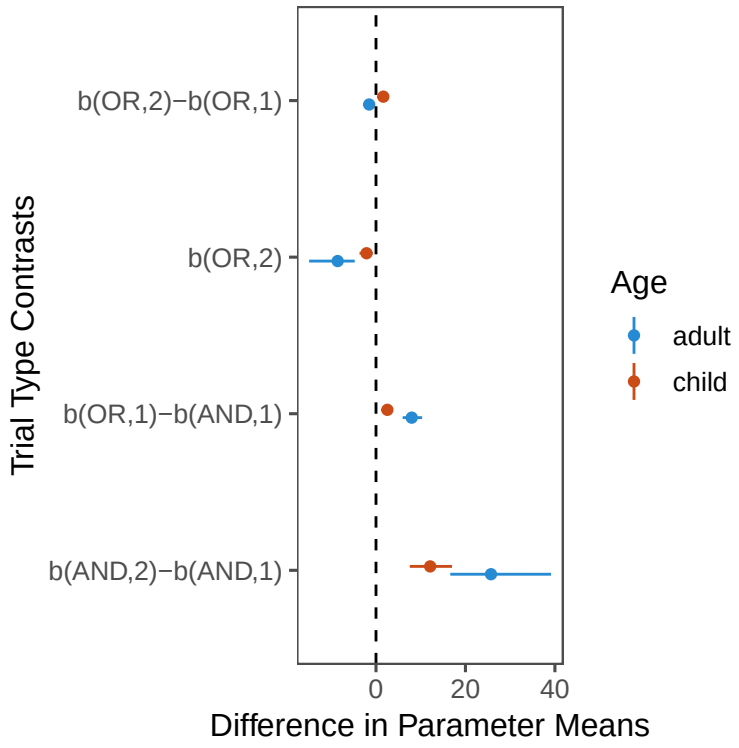

Figure 11. Coefficients capturing the relevant comparisons across conditions in 3AFC judgments in Study 1 and 2. In naming the coefficients like b(OR,2), OR/AND represents the connective used and the number 1/2 represents the number of animals on the card. Error bars represent 99% regions of highest posterior density.

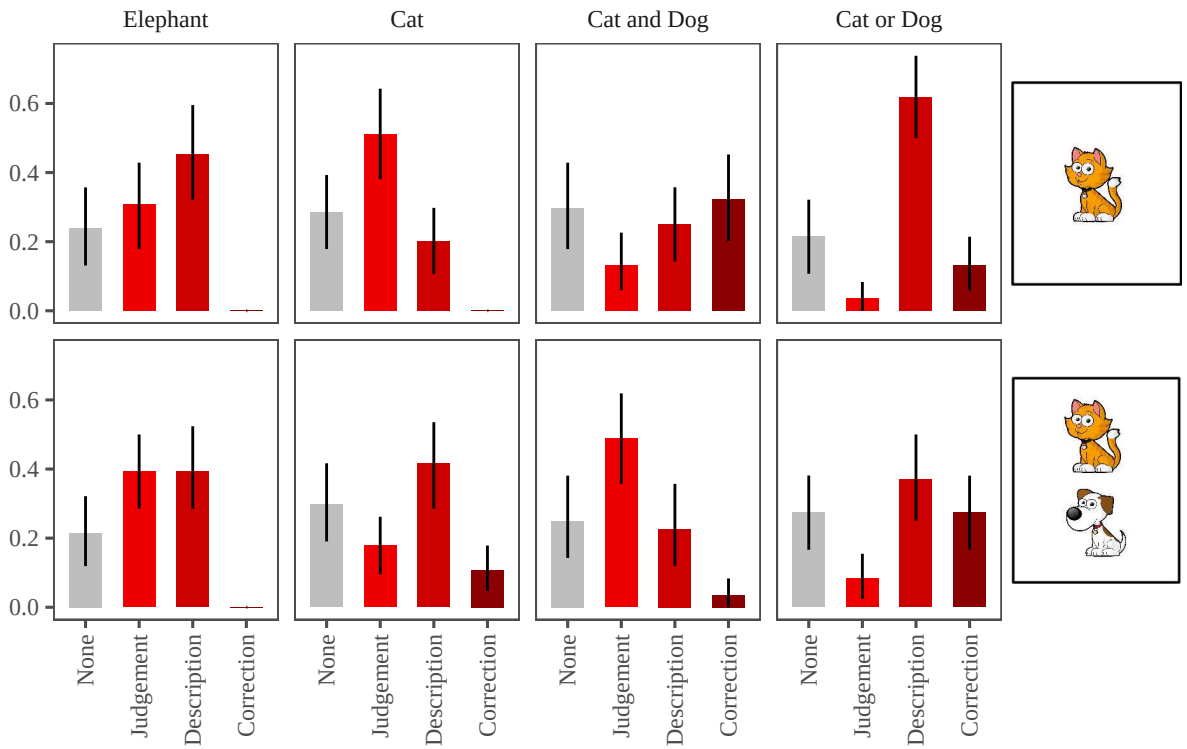

Figure 12. Children's open-ended Feedback. Error bars represent 95% confidence intervals.

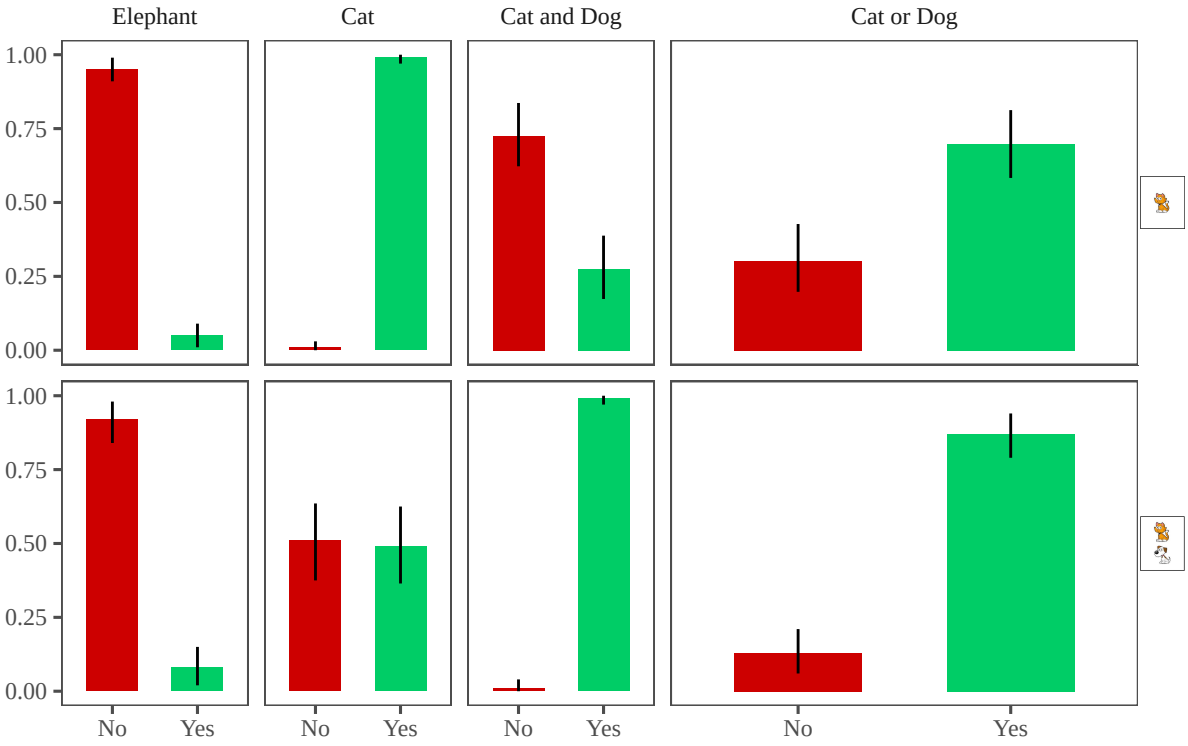

Figure 13. Children's binary truth value judgments.

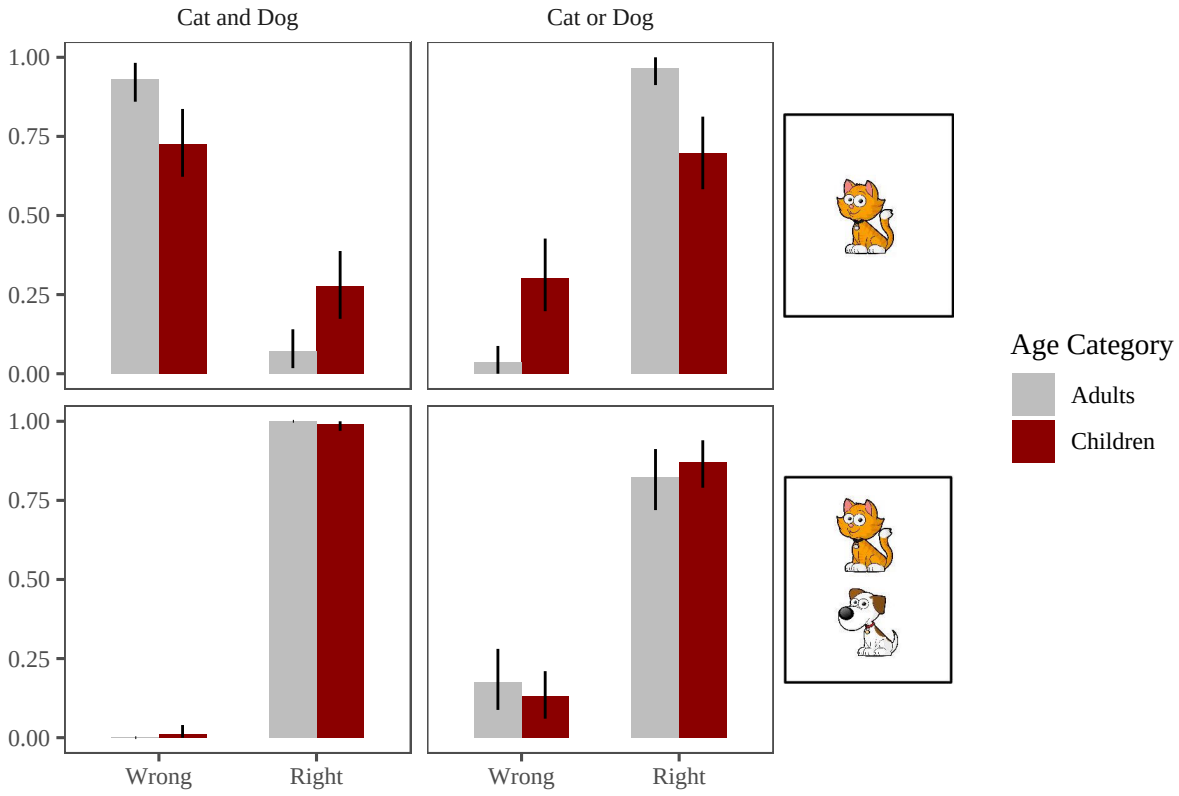

Figure 14. The comparison of the 2AFC judgment task for conjunction and disjunction trials in adults (study 1) and children (study 3).

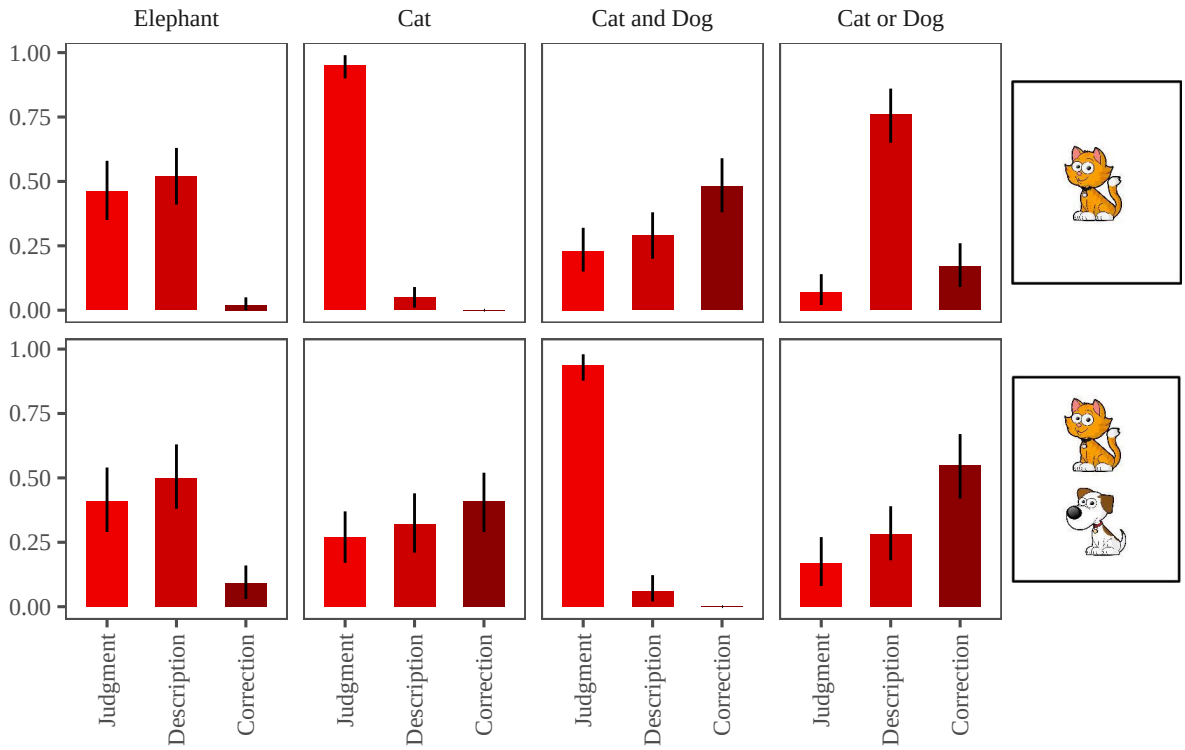

Figure 15. Children's Open-ended Feedback in Study 3. Error bars represent 95% confidence intervals.

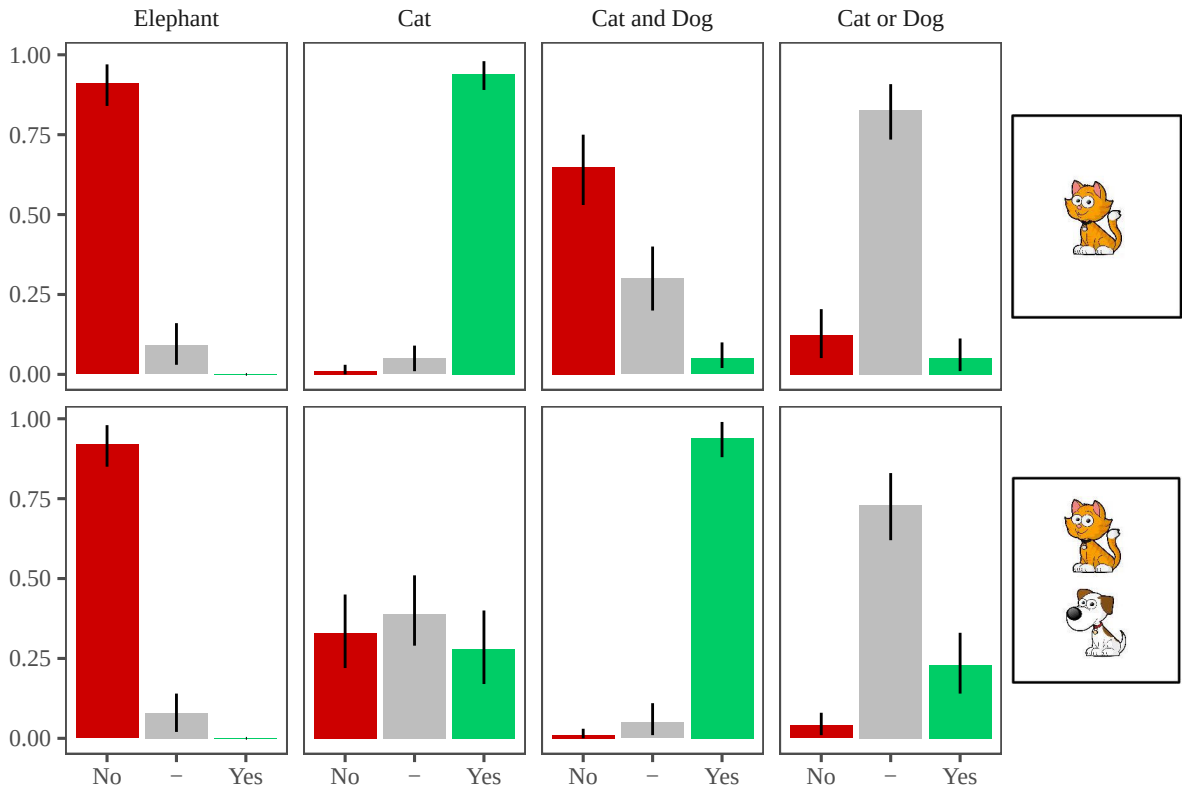

Figure 16. Children's open-ended feedback to the puppet's guesses. The x-axis shows whether children spontaneously provided a yes (green), no (red), or other response (grey).

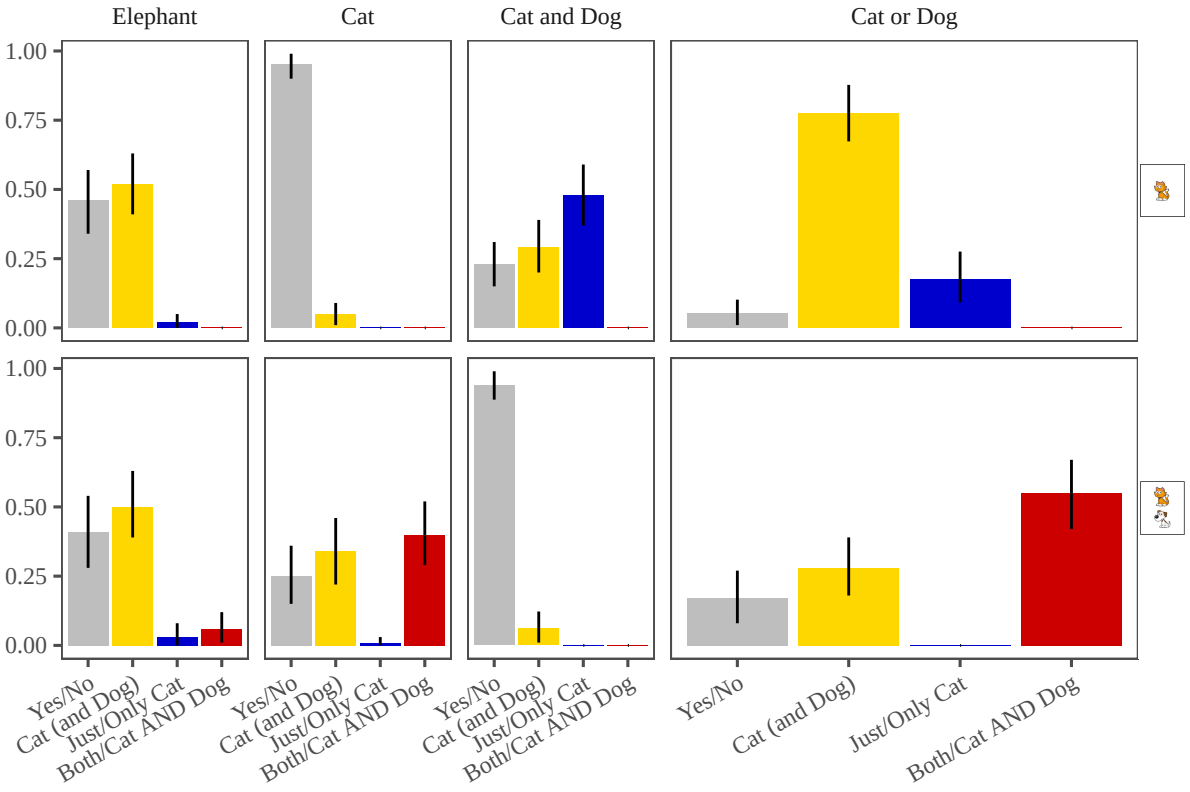

Figure 17. Children's feedback categories in disjunction trials.

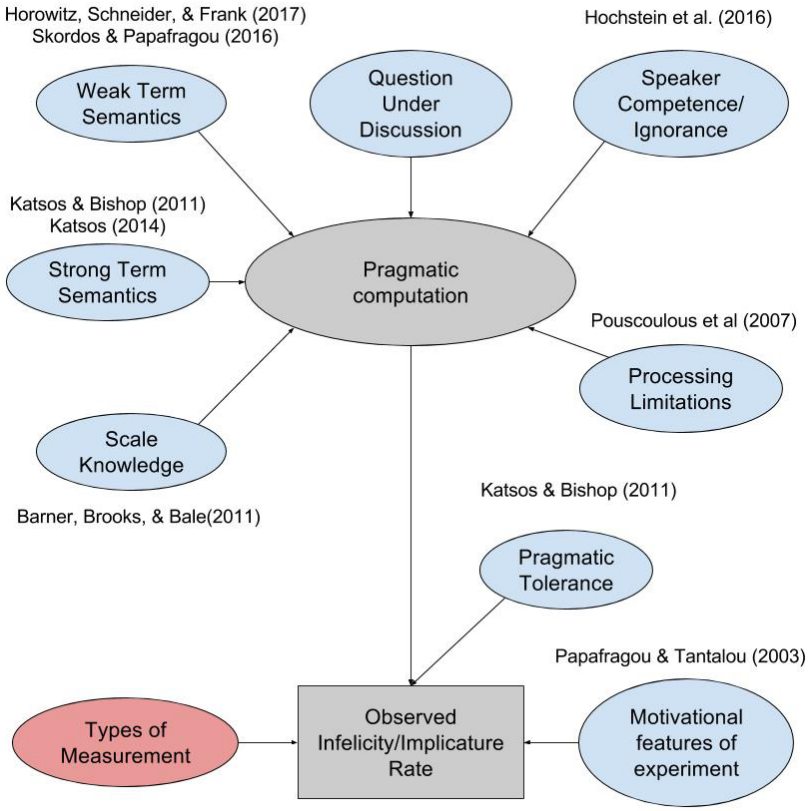

Figure 18. Factors that could affect pragmatic computations and the estimates of these computations in the experimental settings