Group Construal and Preschoolers' Generation of Scalar Implicatures

André Eliatamby (aeliatamby@gradcenter.cuny.edu) and
Sandeep Prasada (sprasada@hunter.cuny.edu)

Abstract

Unlike adults, 4- and 5-year-olds do not interpret *some* as meaning *some but not all*. Two experiments showed that 4- and 5-year-olds will interpret *some* to mean *some but not all* when given cues to construe multiple entities as forming a group. Experiment 1 showed that large numerosities, which are likely to be perceived as ensembles and represented as a group, facilitated children's interpretation of *some* to mean *some but not all*. Experiment 2 showed that describing a small numerosity with the collective noun *family* facilitated children's interpretation of *some* to mean *some but not all.* The results suggest the mechanisms that generate scalar implicatures are in place by age 4, and point to the role of conceptual and morphosyntactic factors in implicature generation.

1. Introduction

A key aspect of human language is the distinction between what is meant and what is said. A speaker who utters (1a) is typically understood to mean (1b):

1. a. Some of the kids were selling cookies.

   b. Some but not all of the kids were selling cookies.

   c. All of the kids were selling cookies

Why does *some* imply *not all* in (1a)? One famous proposal argues that the basic meaning of sentences like (1a) are enriched with *scalar implicatures*, inferences that listeners make about speakers' intended meaning based on the assumption that cooperative speakers are maximally informative (Grice, 1967; Horn, 1972; Levinson, 2000). When a speaker says (1a), listeners infer that the speaker believes that a more informative alternative sentence containing *all,* like (1c), is false. Adding the falsity of (1c) to the meaning of (1a) yields (1b).

Scalar implicatures have been of interest to many cognitive scientists because generating them depends on the coordination of many distinct types of knowledge, including the ability to link together lexical items on the basis of informativity, the ability to reason about speaker intention, and an understanding of the communicative expectations of the discourse.

Preschool-aged children are known to have difficulty generating implicatures (Noveck 2001; Papafragou & Musolino, 2003; Papafragou & Tantalou, 2004; Guasti et al., 2005; Huang & Snedeker, 2009; Katsos & Bishop, 2011; Foppolo et al., 2012; Skordos & Papafragou, 2016; Horowitz et al., 2018; Foppolo et al., 2021; Rees et al., 2023). Many studies have focused on *some,* showing that children accept uses of *some* in situations where *all* is true (Pouscolous et al., 2007, Barner et al., 2011, Foppolo et al., 2012, Skordos & Papafragou, 2016). For example, in a

scenario where there are 4 smurfs and all of them are on a boat, 5-year-olds find "*some of the smurfs are on the boat*" to be an acceptable description of the scenario (Foppolo et al., 2012). Children's divergence from adult performance has been attributed to a variety of sources: a lack of cognitive resources (Pouscolous et al., 2007; Wang, 2023), being pragmatically tolerant (Katsos & Bishop, 2011), having difficulty calling up the lexical item needed to form the alternative sentence (Barner et al., 2011; Horowitz et al., 2018), difficulty tracking a speaker's mental states (Papafragou et al., 2018), and having difficulty determining what is relevant (Skordos & Papafragou, 2016), amongst others.

In this paper, we investigate a factor that hasn't previously been studied: the crucial role that the construal of the situation plays in generating implicatures. For example, to generate the inference that *some* implies *not all,* the entities involved need to be construed as forming a group, rather than as merely a multiplicity of individuals*,* and that quantification must operate relative to that group. For convenience, we will call this the "construal hypothesis." As we discuss below, uses of *some* with plural nouns in adult language don't always represent multiple entities as forming a particular group. This allows for the possibility that children interpret *some* as asserting the existence of a plurality of individuals, potentially of a small cardinality. The experiments we report here examine whether linguistic and non-linguistic cues to group construal facilitate implicature generation in preschoolers.

In adult language, phrases using *some* have at least two distinct interpretations, only one of which contrasts with *all.* What we call the "group relative*"* interpretation, shown in (1a), interprets *some* relative to a group of entities, and asserts that the described property applies to

members of that group (Reed, 1991; Degen, 2015).[1] In (1a), the relevant group is denoted by *the kids,* and (1a) asserts that a subset of that set of children were selling cookies. In English, this interpretation is unambiguously signaled by the partitive morphosyntax, "*of the X."* The other use, what we call *"*multiplicity asserting," simply asserts the existence of a multiplicity, as shown in (2):

2. Some kids were selling cookies at the carwash.

If spoken without emphasis on *some*, and if there are no salient, familiar, or mentioned children in the discourse, (2) asserts the existence of a multiplicity of children that were selling cookies and introduces them into the discourse, much like an indefinite determiner (Heim, 1983; Diesing, 1992). Importantly, only the group relative use of *some* contrasts with *all*, as the following contrasts show. In (3), *some* appears in a partitive phrase, and contrasts with *all,* as the felicity of the "in fact" continuation reveals:

3. Some of the kids were selling cookies. In fact all of the kids were selling cookies.

In (4), if *some* is spoken in the same way as in (2), the contrast with *all* is infelicitous:

4. *Context: There are no kids salient or previously mentioned in the current discourse*
   Some kids were selling cookies at the carwash. #In fact, all kids were selling cookies at the carwash.

On the other hand, the multiplicity asserting use of *some* seems to contrast with *many* or *a lot,* as in (5), suggesting that this use of *some* may imply a small numerosity:

[1] This interpretation has also been called "proper partitivity" (Selkirk 1977). We avoid this term because "partitive" is also used to describe the morphosyntax of the determiner phrase *some* (and other lexical items) appear in.

5. *Context: There are no kids salient or previously mentioned in the current discourse*

   Some kids were selling cookies at the carwash. In fact, a lot of kids were selling cookies at the carwash.

These intuitions are supported by experimental studies showing that for adult English speakers, *some* more strongly implies *not all* when used in a partitive frame than a non-partitive frame (Degen, 2015; Degen & Tannenhaus, 2015). Degen & Tannenhaus (2015) for example, found that in an acceptability rating task, English-speaking adults rated uses of *some* with partitive syntax lower than uses of *some* without partitive syntax in scenarios where *all* was true, and that *some* is rated highest when used to describe between 3 and 6 entities.

Languages have a variety of ways of distinguishing between group relative and multiplicity asserting interpretations of *some.* In English, partitive morphosyntax unambiguously marks the group relative interpretation, while the non-partitive form, *some Xs,* is ambiguous between the multiplicity asserting and the group relative interpretation in the absence of prosodic or contextual cues. Italian is similar (Foppolo, 2012; Giusti, 2021). French, Dutch and Romanian seem to have distinct lexical items but can also make use of partitive morphosyntax (Pouscolous et al., 2007; Corblin,et al., 2014; de Hoop, 1995, Davatz & Stark, 2024). Spanish (Martí, 2008) and Polish (Tomasz Zyglewicz, p.c.) use distinct lexical items. German can use the same lexical item and rely on context for disambiguation (Benz & Salfner, 2011). Children must learn how their particular language distinguishes between these two uses of *some*, and until they acquire that linguistic knowledge, they will have to rely on non-linguistic and contextual cues that indicate whether a multiplicity should be construed as a group or as a number of individuals.

All of this suggests that children's lack of implicatures may be due to a lack of knowledge of the semantics of the partitive and/or their lacking the understanding that *some* has a group relative interpretation.[2]

We investigate the hypothesis that English-learning children are capable of generating scalar implicatures when given clear nonlinguistic and linguistic reasons to treat multiple entities as a group. The construal hypothesis predicts that children will be able to generate *not all* implicatures with *some* "out-of-the-blue" (in the absence of an explicit linguistic or pictorial contrast) if they are given reason to treat the entities being presented as forming a group. In the absence of such reasons, children will interpret *some* as being compatible with *all* and possibly also to imply a small cardinality, which would be consistent with them interpreting *some* to assert the existence of a multiplicity.

Relation of the Construal Hypothesis to Previous Findings

While young children often interpret *some* as being compatible with *all*, some studies have found that under certain circumstances, children are able to generate *not all* implicatures or are sensitive to the infelicity of using *some* when *all* is true. One important study by Skordos & Papafragou (2016) highlights the role of making quantity relevant in facilitating implicature generation. In Skordos and Papafragou's (2016) Experiment 2, 5-year-olds assessed true and

[2] We are admittedly simplifying greatly in our description of the cross-linguistic data. Many languages distinguish properties beyond just the group relative and multiplicity asserting interpretations we are focusing on (see, for example, Seržant's 2021 typological overview). For example, whether quantification is over kinds (Chierchia 1998) or whether the items being quantified share a certain property (de Hoop 1995), must also be acquired in languages that make such distinctions. We focus on representing a group versus a collection of unrelated individuals because non-linguistic evidence suggests that these distinctions are particularly salient during development (as we discuss below). However, the variation found in the form-meaning mapping of these constructions increases the learning challenge for the child, and makes it less likely that their literal interpretation of *some* is not adult-like, as previous studies have generally assumed.

false sentences using *all* before assessing sentences using *some*. In the "quantity" condition, children assessed uses of *all* that were false due to an incorrect quantity – e.g., because only 3 out of the 4 entities had bones – while in the "object" condition, children assessed sentences where *all* was false due to an incorrect property – e.g., because the four entities had umbrellas rather than bones. Only the children who assessed sentences where *all* was false due to quantity subsequently generated implicatures when interpreting *some*. In Experiment 3, a third group of children first assessed true and false sentences with *none,* and these children also generated implicatures. This is notable, because *none* is not used in the generation of *not all* implicatures, since *none* contradicts *some.* Skordos and Papafragou (2016) propose that children need assistance determining that quantity is the relevant dimension of assessment for the task. Quantity is highlighted when *all* or *none* is false due to an incorrect quantity. Similar behaviour with Italian-speaking 5-year-olds was observed in Foppolo et al.'s (2012) Experiment 6, where like Skordos & Papafragou's "quantity" condition, children who first assessed uses of *all* that were false before assessing statements using *some* generated *not all* implicatures.

Highlighting the relevance of quantity for interpreting *all* is consistent with highlighting the relevance of **proportion**, a type of quantity. Thus, it may be that the manipulations in Skordos and Papafragou's (2016) experiments highlighted the relevance of proportion (as opposed to numerosity), and that this facilitated the group relative interpretation of *some.* Proportions may recruit group representations because (at least in countable contexts) they involve representing multiple entities as forming a part of a larger whole. For example, consider a sentence like *all of the blickets have bones* from Skordos and Papafragou's (2016) Experiment 2. In scenes where only 3 out of 4 blickets have bones, determining this sentence is false requires understanding that only a proper subset of entities have the described item, and representing a

proper subset of the entities requires representing multiple entities as forming a part of a larger whole. Similarly, attending to the proportion of blickets with bones can help determine whether *none of the blickets have bones* is false. In contrast, when assessing *all of the blickets have bones* when they all have umbrellas, children can reject the sentence by simply noting the absence of any bones. The blickets do not have to be represented as a group to verify that the sentence is false.

In another sort of study, Pouscolous et al. (2007) found that children seem to generate *not all* implicatures in an act-out tasks. 4-to-7-year-old French-speaking children were asked to make changes to objects based on statements like, *"I would like some boxes to contain a token."* In situations where initially all boxes contained a token, they found that even the 4-year-olds removed a token from a box at rates greater than chance, suggesting they interpreted the statement to mean the experimenter wanted some but not all boxes to contain tokens. Notably, however, their task was ordered in such a way that children had to act on instructions using *all* and *no* that required attending to quantity (e.g. adding tokens to satisfy the instruction *"I would like all boxes to contain a token*") before acting on instructions using *some* in scenarios where all boxes contained a token. Thus, the study may have provided the sort of linguistically contrastive support for attending to quantity (or, if we are correct, more specifically proportion rather than simply numerosity) like in Skordos & Papafragou (2016).

Other experiments where children successfully generate implicatures with *some* involve scene selection tasks where children must select the scene or scenario best described with a sentence using *some* (e.g., Katsos & Bishop, 2011, Papafragou et al., 2018, Rees et al., 2023). For example, in Katsos and Bishop's (2011) Experiment 3, 5-year-olds were presented with sentences like "*Mr. Caveman took some of the carrots"* and four scenes. One scene had Mr.

Caveman taking only some of the carrots in question, another had Mr. Caveman taking all of them, while the other two had distractor scenes with Mr. Caveman taking no carrots. When asked which scene the sentence described, 5-year-olds were successful in selecting the scene where only some carrots were taken, evidence that they can interpret *some* to mean *some but not all.* In these tasks, the visual contrast may have highlighted differences in proportion, rather than a mere numerosity, facilitating the group relative interpretation of *some,* and thus facilitating implicature.

In addition to these studies, there are others that do not directly show evidence of implicature generation, but suggest that children do not find *some* to be the best description of scenarios where *all* is true. In ternary judgement tasks, for example, children are asked to assess descriptions of scenes using a ternary scale, where the middle option of the scale is in between "full acceptance" and "full rejection" (Katsos & Bishop, 2011; Tieu et al., 2019). For example, Katsos & Bishop's (2011) Experiment 2 found that when assessing descriptions using *some* where *all* is true, 5-year-olds pattern like adults in choosing the middle option. While such results are consistent with children being sensitive to underinformative uses of *some,* they leave open the question of what allows children to successfully generate scalar implicatures in the scenarios they have been shown to do so. The prediction of the construal hypothesis is that cues that encourage representing the entities as forming a group are an important factor that lead to such success. As such, the results of ternary judgement tasks are consistent with the construal hypothesis though they don't directly address it.[3]

[3] The results of ternary judgement tasks are also consistent with children being sensitive to the ambiguity of *some*, i.e., sensitive to the fact that phrases headed by *some* have both the group relative and multiplicity asserting interpretations. Thus, they do not necessarily reveal a sensitivity to underinformativity. Further research is needed to understand what capacities the ternary tasks track in this case.

Finally, the construal hypothesis is also consistent with the many studies where children do not generate implicature with *some* that provide no reason to treat the multiple entities as a group other than the partitive syntax (e.g. Skordos & Papafragou, 2016; Horowitz et al., 2018; Foppolo et al., 2012; Papafragou & Tantalou, 2004; Huang & Snedeker, 2009, Katsos & Bishop, 2011; Pouscoulous et al., 2007; Novek, 2011; *inter alia*). Only one study with young children explicitly compared implicature rates between partitive and non-partitive forms. Foppolo et al. (2012) found no difference in 5-year-old Italian-speaker's ability to generate implicatures with the partitive (*alcuni dei*) and the bare existential *qualche.* Adults, unlike young children, can make use of the morphosyntactic properties of their language to determine which interpretation of *some* is appropriate.

In sum, though we cannot know that children's successful generation of *not all* implicatures in previous research involved group construals, many of the experimental conditions in which that happened, plausibly encouraged group construals by linguistically or visually highlighting contrasting proportions.

Non-linguistic Influences on Group Construal

Nonlinguistic research shows that children and adults represent small and large numerosities differently (Feigenson, et al., 2002; Feigenson & Carey, 2005; Choo & Franconeri, 2014). Small numerosities seem to be represented as individuals via object files (Feigenson & Carey, 2005; Le Corre & Carey, 2007; Carey, 2009), while larger numerosities elicit ensemble representations that collapse over individuals and facilitate the assessment of group-level properties like average size and average color (Ariely, 2001; Alvarez, 2011, Maule et al., 2014). Psycholinguistic studies have shown that the interpretation of universal quantifiers like *each* and *every* make use of different construals of multiplicities, whereby *each* construes the multiplicity

as consisting of independent individuals, and *every* makes use of a construal where the multiplicity is an ensemble collection (Knowlton, 2021; Knowlton, Halberda, Pietroski & Lidz, 2023). For example, Knowlton, Halberda, Pietroski & Lidz (2023) show that adults are better at detecting the change in colour of one particular element of an array, a property of an individual, when the array is described with *each* compared to when it is described with *every.* With children, Knowlton (2021) showed that 5 to 8-year-olds are better at recalling the centre of a collection of circles, a property of the ensemble, when they've been primed with *every* compared to when they are primed with *each.*

Other non-linguistic research has shown that children have difficulty attending to proportional information, particularly in contexts with countable, discrete units (Boyer et al., 2008, Jeong et al., 2007; Hurst & Cordes, 2018). One linguistic reflection of this is in children's interpretation of *most,* where children are found to attend to numerical information over proportional information (Sullivan et al., 2018; Hurst & Levine, 2022). This may contribute to children's difficulty in generating scalar implicatures which we hypothesize requires a group relative interpretation of *some.*

Most developmental studies investigating *some* present children with small numerosities – between three and six entities to assess (e.g. Skordos & Papafragou, 2016; Horowitz et al., 2018; Foppolo et al., 2012; Papafragou & Tantalou, 2004; Huang & Snedeker, 2009, Katsos & Bishop, 2011; Pouscoulous et al., 2007; *inter alia*). Thus, children in these studies may have been interpreting *some* in contexts which provide a strong signal to attend to individuals over ensembles and to attend to cardinalities over proportional information. The construal hypothesis predicts that in such situations, children are less likely to generate implicatures because they are not given reason to treat the entities being presented as forming a group.

The experiments in this paper explicitly test the role of group construal in the interpretation of *some*. We do this by manipulating the number of items children see in different scenarios (Experiment 1) and by providing a linguistic cue to treat the items as forming a group (Experiment 2). We use a design where trials with *some* are presented before trials with *all* (as in Skordos & Papafragou's *Some-First* condition in Experiment 1) so that children must generate implicatures "out of the blue" – i.e., in the absence of contrastive information made available via explicit uses of *all*.

Under the assumption that (i) children have nonlinguistic biases to represent small numerosities as individuals via object-files, (ii) that large numerosities elicit ensemble representations that may encourage group construals, and (iii) that preschool-aged children have not mastered the semantics of the partitive construction, we make the following predictions:

(A) Children will construe a large number of items (e.g. 20) as a group, and a small number of items (e.g. 4) as simply a multiplicity of individuals. Children should be more likely to reject *some* to describe a situation where all items have a property when the numerosity is large (e.g., where 20/20 items have the property) than when the numerosity is small (where 4/4 items have the property). In contrast, adults, who understand the meaning of the partitive syntax, will reject the use of *some* when *all* is true equally in both large and small numerosity contexts.

(B) Children who hear the group denoting noun *family* used to talk about a small number of items (e.g. 4) will construe the items as a group and thus be more likely to reject the use of *some* when it describes all the items (e.g. 4/4) than children who do not hear the items described using the group-denoting noun.

Our experimental design also allows us to test whether children interpret *some* to simply mean a small numerosity. Those data bear on whether children might have been accepting the use of *all* in previous studies because they considered the items presented to constitute a small numerosity. If children think *some* can only be used to talk about a small numerosity, we predict:

(C) Children will reject the use of *some* in the large numerosity context, regardless of whether it describes both a proper subset (15/20) of items or all the items (20/20), while accepting its use in both small numerosity contexts (e.g. when 3/4 or 4/4 items have the mentioned property).

Experiment 1 tested predictions (A) and (C). Experiment 2 tested prediction (B).

2. Experiment 1

Experiment 1 tested predictions A and C using a Truth Value Judgement Task (Crain & Thornton, (1998) by varying the total set size of the entities being presented. *Trial type* was a within subject variable with two levels (*some-subset*, *some-full*), *numerosity* was a between subject variable with two levels (*small*, *large*), and *age* was a between subjects variable with three levels (*age 4*, *age 5*, *adults*). We tested 4- and 5-year-olds because this seems to be the age range during which the robust generation of implicatures is under development. Many studies have shown that both 4- and 5-year-olds struggle to generate *not all* inferences with *some* (e.g., Papafragou and Musolino, 2003; Barner et al., 2011; Skordos & Papafragou, 2016; Horowitz et al., 2018). While age 4 is the earliest age at which children have been shown to be successful (Papafragou et al., 2018), studies that explicitly compared both age groups find 5-year-olds to be more capable (Papafragou et al., 2018, Horowitz et al., 2018).

If prediction A is correct, we expect a three way interaction between trial type, numerosity, and age, such that the 4- and 5-year-olds show a two-way interaction between trial type and numerosity, such that in the large numerosity condition, they reject *some* more often in full trials than the partial trials but do not differ on full and partial trials in the small numerosity condition. It's also possible that we will see this effect with 5-year-olds but not 4-year-olds, since as noted above, 5-year-olds have been shown to be more capable of generating implicatures than 4-year-olds. Adults in contrast are expected to reject *some* more often in full trials than partial trials in both numerosity conditions.

If prediction C is correct, we expect children to show a main effect of numerosity, no effects of trial type, and no interaction between trial type and numerosity. I.e., they reject the use of *some* in large numerosity contexts independent of trial type.

Participants

We recruited and tested 155 English-dominant, typically developing 4- to 5-year-olds (M=4.99, min= 3.82, max=5.99). This included 80 4-year-olds (M=4.52, min=3.82, max=4.99) and 75 5-year-olds (M=5.51, min=5.02,max=5.99). Almost all children spoke U.S. English; two spoke Australian English. We recruited participants until there were at least 25 children in each age group and in each scene size condition who could be included for analysis according to predetermined inclusion/exclusion criteria (see *Procedure* section). Participants were recruited through internet advertising and word of mouth. Participants were excluded if they failed the control trials (n=44), because of parental interference (n=2), and because of experimenter error (n=1).[4] In total 108 children (50 4-year-olds and 58 5-year-olds) were included for analysis. We

[4] We consider implications of the high exclusion rate in the results section.

also ran 60 adults residing in the United States and aged between 22 and 63 that had been screened as having English as their first language, being dominant in English, and not having any language, speech or cognitive impairments.

Stimuli and Design

Stimuli consisted of scenes containing common animals with objects attached to them paired with pre-recorded sentences. *Small numerosity* scenes presented 4 animals, while *large numerosity* scenes presented 20 animals. In *subset* configurations, 3/4 and 15/20 animals possessed an object. In *full* configurations, 4/4 and 20/20 animals possessed an object (Figure 1). Sentences were assertions recorded by a female speaker involving the quantifier *some* or *all* in subject position, using partitive syntax, and with a singular noun phrase in object position: e.g., "some of the dogs have a balloon"; "all of the fishes have a crown." Pictures and sentences were combined to create 4 trial types: *some-full, some-subset, all-full,* and *all-subset.*

An individual scene featured only one type of animal and one type of object. Overall, 16 different animals and four objects were used yielding 64 scenes for each numerosity condition. These were divided into four experimental lists with 16 trials, four of each trial type (see *Appendix*). The lists counterbalanced animals and objects with each trial type, such that each animal appeared only once in each list, and across the 4 lists, each animal occurred once for each trial type. Each object appeared 4 times in each list, once for each trial type. Furthermore, across the 4 lists, an animal-object pairing occurred only once. Lists were identical for large and small scenes; the only difference was whether 4 or 20 animals were presented.

Four training stimuli for each list were created showing a single animal with one object, and paired with a prerecorded sentence like *"the horse has a hat."* (Figure 2). Two of the

training stimuli were true, one was false because the object mentioned was incorrect, and one was false because the animal mentioned was incorrect.

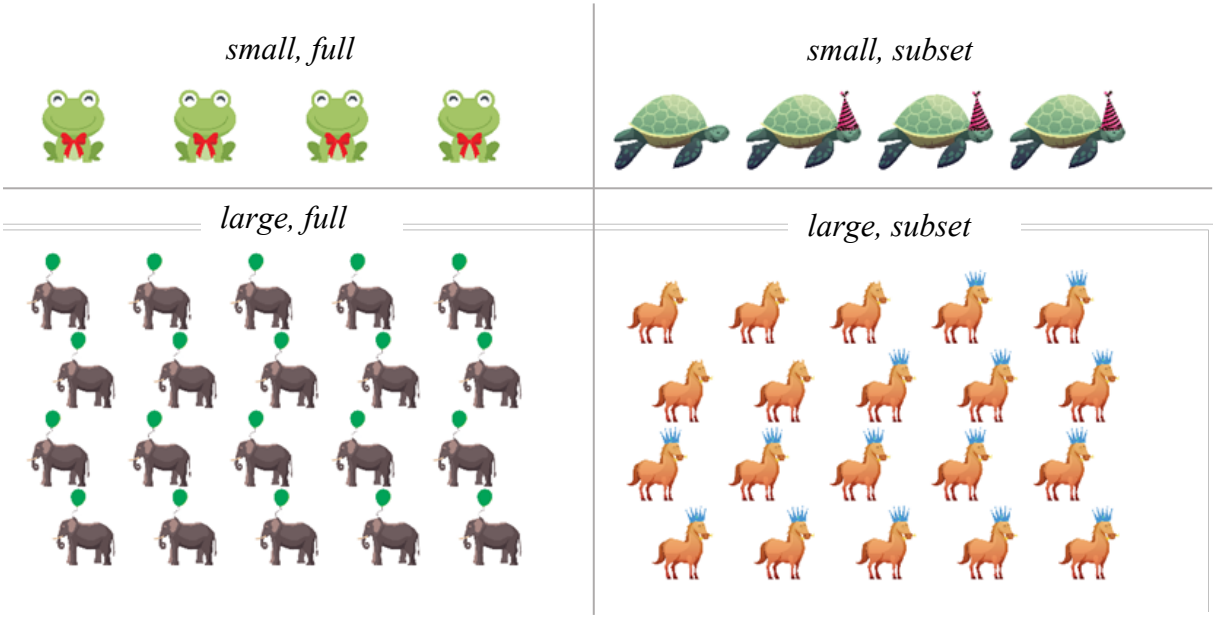

Figure 1: Large/Small scenes in full/subset configurations. Trials paired these scenes with sentences involving "some" or "all." E.g. "some of the elephants have balloons" and "all of the turtles have hats."

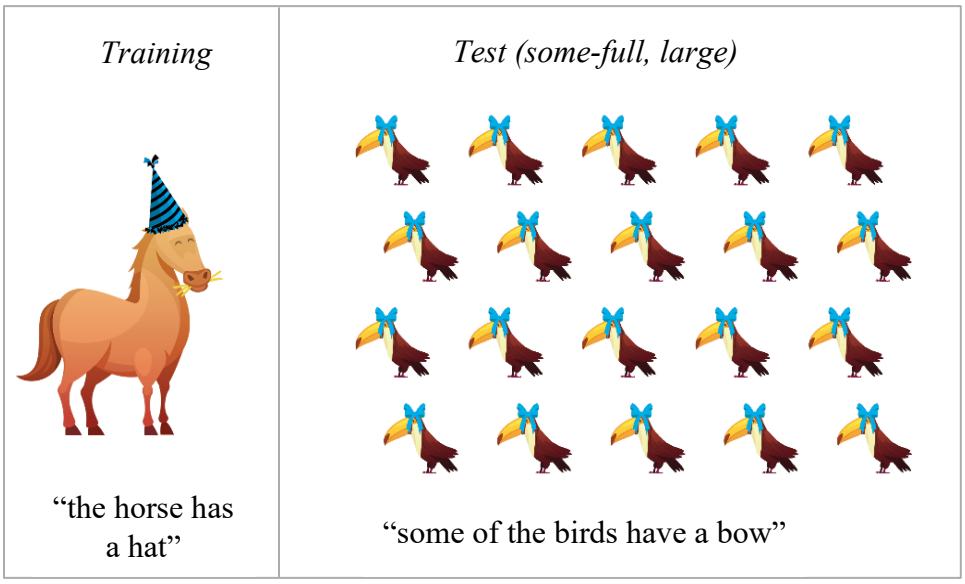

Figure 2: Examples of training and test trials.

Importantly, the order of the trials was blocked to avoid the effects that occur when children assess sentences with *all* prior to assessing underinformative statements using *some* (Skordos & Papafragou, 2016; Foppolo et al., 2012). In the first block, children assessed eight trials involving *some* (i.e. 4 *some-full* trials and 4 *some-subset* trials, pseudo-randomized). In the second block children assessed eight trials involving *all* (i.e. 4 *all-full* trials and 4 *all-subset* trials, pseudo-randomized). Block ordering was consistent for all participants. Thus, children cannot make use of linguistic or visual contrast between *some* and *all* trials to generate implicatures

Procedure

Participants were pseudo-randomly assigned to conditions with the aim of ensuring balance across conditions. For child participants, the study was conducted using PsychoPy

remotely over Zoom. A single experimenter met with the child and a guardian over Zoom. Guardians were asked to not assist their child in their responses. After introductions and familiarization, children were introduced to a puppet. Participants were told that the puppet was going to talk about pictures, and that they needed to tell the puppet if what it said was right or wrong. Participants were also asked to tell the puppet why it was wrong if they thought it was wrong. Children then did the 4 training trials and 16 test trials. Participants were **not** excluded if they answered the training trials incorrectly. In these cases, they were asked to explain their response and then provided with the correct response and a reason why that response was correct. The experimenter then proceeded to the test phase. In the test phase, the scenes were presented first, with the description playing after a 1-second delay. After the description played, children were asked to repeat what the puppet said, and then asked whether the puppet was right or wrong. If they said that the puppet was wrong, they were asked for a justification, which the experimenter noted. Some children repeated the sentence automatically and volunteered the assessment and justification without prompting. For adult participants, the experiment was administered asynchronously using the online platform, Pavlovia (https://pavlovia.org/). The adults also filled out a screening survey about the languages they spoke and basic demographic information.

Exclusion Criteria

We decided in advance to exclude participants from our analysis if they did not display an understanding of the basic meanings of *some* and *all.* Participants were excluded if their responses followed any one of the below patterns:

- they accepted three or more *all-partial* trials.

- they rejected three or more *all-full* trials.
- they rejected three or more *some-partial* trials, while also accepting three or more *some-full* trials.

This criterion led to 44 children being excluded from analysis (30 four-year-olds, and 14 five-year olds). No adults met the exclusion criteria.

Results

Figure 3 shows the mean proportion of trials where *some* was accepted across the different conditions. We analyzed the results using a mixed effect logistic regression implemented in *R* using the *lme4* package. Recall that the prediction A – i.e. the construal hypothesis – predicts a three-way interaction between trial type, numerosity, and age, such that unlike adults, 4-year-olds and 5-year-olds will reject *some* more often in full trials than the subset trials in the large numerosity condition but not differ on full and subset trials in the small numerosity condition (i.e., that for 4-year-olds and 5-year-olds only, we will find a two-way interaction between trial type and numerosity).

Our initial model estimated participant response (*accept* or *reject*) as a function of trial type (*subset*, *full*), numerosity condition (*small*, *large*), age category (*age 4*, *age 5*, *adult*) and all two and three way interactions. The random effects structure included random intercepts for participants and items, with trial type specified as random slopes against the participant intercepts.[5] We compared the initial model to a reduced model without the three-way-interaction using a likelihood ratio test.

[5] initial model formula: response ~ 1 + (trial.type * age.bucket * condition) + (1 + trial.type | participant) + (1 | item)

The initial model with a three-way interaction had a significantly better model fit than the reduced model ($\chi^2(2) = 12.3$, $p < 0.003$). To follow up the three-way interaction, we performed FDR-adjusted Wald tests using the *emmeans* package to test for significant two-way interaction between trial type and numerosity condition for each age group. The interaction predicted by prediction A was found for the 4-year-olds (β=-14.8, $p < .001$), but surprisingly not for the 5-year-olds (β=5.0, $p = .21$), who behaved in an adult-like manner by rejecting *some* on full trials more often than on subset trials (β=-18.4, $p < .0001$) and showed no effect of numerosity (β=.19, $p = .93$). Adults behaved as predicted, showing no effect of the interaction between trial type and numerosity (β=2.4, $p = .56$), no effect of numerosity (β=-.85, $p = .66$), and rejecting *some* more often in full trials than in subset trials (β=-17.1, $p < .0001$).

To examine age group differences in *some-full* trials, we conducted FDR-adjusted pairwise comparisons of all three groups' response to the full trials in the large numerosity condition and small numerosity condition. In the large condition, 4-year-olds, 5-year-olds, and adults did not differ from each other in their rejection of full trials (age 4 vs. age 5: β=-.26, $p = .9$; age 4 vs. adults: β=-.43 $p = .9$**)**. In the small condition, 4-year-olds were more likely to accept the full trials than both the 5-year-olds (β=8.1, $p < .002$**)** and the adults (β=7.0, $p < .003$**)**. 5-year-olds and adults did not differ in either numerosity condition **(**large: β=-.17, $p = .9$; small:β=-.1.2, $p = .62$).

Prediction C – that *some* is interpreted to mean a small numerosity – was not supported by the aggregate data. 4-year-olds showed an interaction between trial type and numerosity, while 5-year-olds showed a main effect of trial type but no effect of numerosity. As a group, children do not interpret *some* to mean a small numerosity. However, there was a small number of participants (6 children and 3 adults) who did provide a response pattern consistent with

prediction C, i.e., they rejected *some* in both the full and subset trials of the large numerosity condition. In their justifications for the rejection, children in this subgroup said things like *"lots of [the birds] have balloons"* or *"a lot of the mice have a hat."* This suggests that the small numerosity interpretation is available for some children, even if it is not the preferred interpretation of children as a group.

Given the relatively large number of 4-year-olds excluded from analysis for failing control trials, one might wonder whether the finding that 4-year-olds reject *some* in full trials more in the large condition than in the small condition is because of a greater participant exclusion rate in the large condition. This does not seem to be the case: 13 out of 38 4-year-olds from the large condition were excluded, while 17 out of 42 were excluded from the small condition. A chi-squared test found no difference in exclusions between these conditions ($\chi^2(1)= 0.12$, $p = 0.73$).

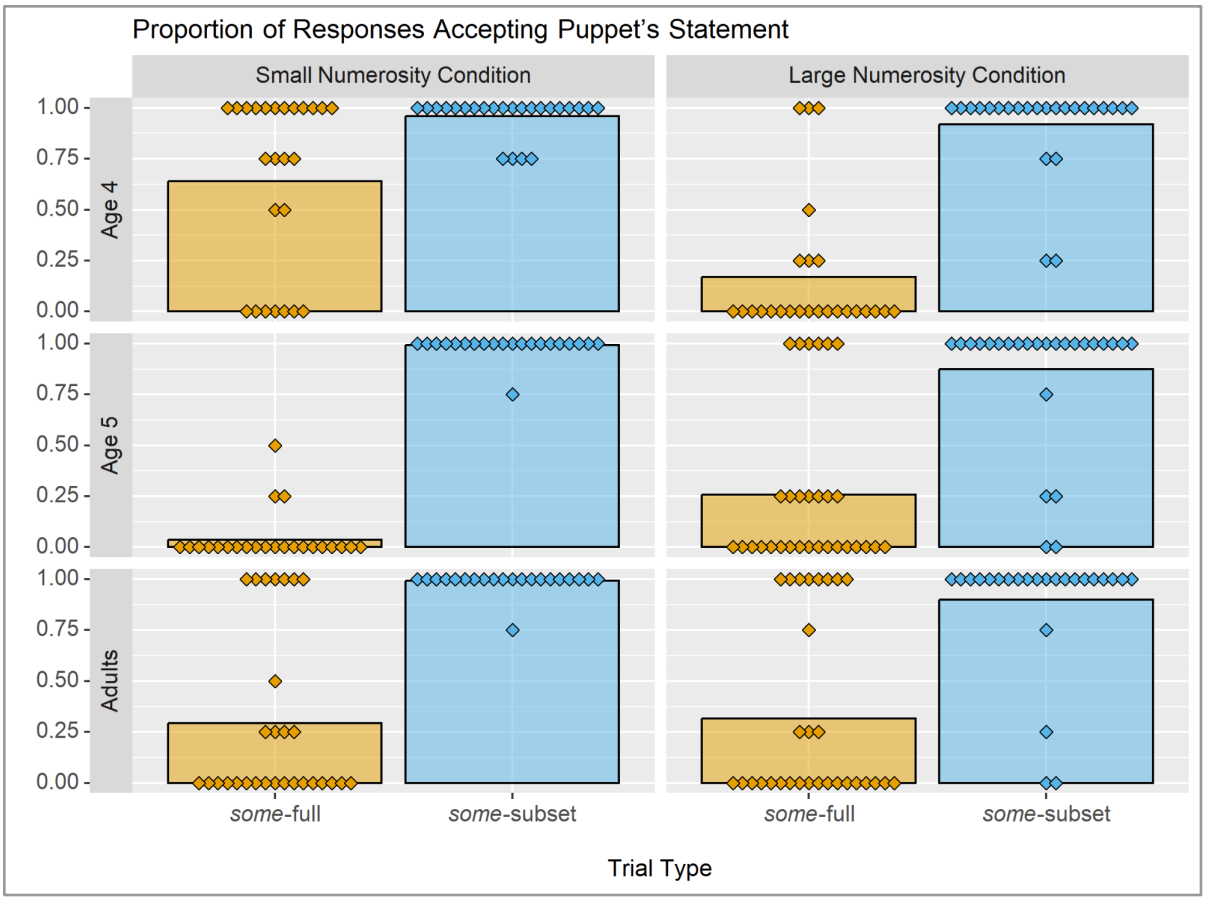

Figure 3: Proportion of responses that accepted the puppet's statement in the small scene condition. Bars show the mean proportion of accepted trials, and diamonds correspond to each participant's individual response. For example, two 4-year-olds accepted the puppet's statement in full trials half the time.

Discussion

Experiment 1 tested the construal hypothesis (prediction A), which predicted that children would be more likely to reject *some* to describe a situation where all items have a property when the numerosity is large (e.g., where 20/20 items have the property) than when the numerosity is small (where 4/4 items have the property). Experiment 1 also tested the hypothesis

that *some* means small quantity (prediction C). If this hypothesis were true, children would reject all uses of *some* in the large condition, independent of whether *some* described all the items or a proper subset of items.

The results of Experiment 1 are consistent with the construal hypothesis, but not the hypothesis that *some* means a small quantity. Children were more likely to interpret *some* to imply *not all* when shown a large number of items than when shown a small number. This suggests they construed the larger numerosity as forming a group, rather than simply a multiplicity of individuals. Children as a group did not reject the use of *some* to talk about large numerosities, suggesting that they do not think *some* refers only to small quantities. However, there was a small subset of participants (6/108 children) who did seem to interpret *some* to imply a small quantity. These children rejected uses of *some* in the large numerosity condition in both the 15/20 and 20/20 configuration. This suggests that even though children generally don't prefer to make this interpretation of *some* (under this particular experimental design at least), it is available.

Our results also show that children as young as 4 years old are able to generate scalar implicatures with *some* "out of the blue," – i.e., in the absence of any explicit contrastive information. Our design did not contrast underinformative uses of *some* with *all;* large numerosities were sufficient to facilitate *not all* implicatures. This is predicted by the construal hypothesis: children will generate implicatures if they are given reason to treat the entities being presented as forming a group.

Surprisingly the 5-year-olds' succeeded in the small scene condition. We aren't sure why the 5-year-olds, but not the 4-year-olds were able to do this. We speculate that our materials may have played a role in cueing the group relative interpretation, but further research will be needed

to determine exactly what caused their success. Experiment 2 focuses on whether the success in the large number condition in Experiment 1 was due to group construal, as hypothesized. Because of this we focus on 4-year-olds in Experiment 2.

3. Experiment 2

According to the construal hypothesis, the key difference between the large and small condition in Experiment 1 is that the large condition facilitated group construals while the small number condition did not. However, the design of Experiment 1 confounds group interpretation and numerosity, making it difficult to conclude that it is group construal, and not large numerosity per se, that facilitates implicature generation.

Experiment 2 seeks to disentangle these two possibilities by using a linguistic cue to group construal while presenting a small number of entities to participants. We use the collective noun *family* to encourage thinking of the entities being presented as belonging to a group – e.g., a horse family. If the construal hypothesis is correct, children encouraged to think of four animals as belonging to the same family will generate more *not all* implicatures when all four animals have the described property than those who do not receive such encouragement.

Method

Participants

We recruited 88 English-dominant, typically developing 4-year-olds (M=4.4, min= 3.6, max=4.99) living in the United States. We recruited participants until there were at least 25 children in each condition who could be included in the analysis according to the inclusion/exclusion criteria used in Experiment 1. Participants were recruited through internet advertising (n = 51) and Children Helping Science (n = 37) (Scott & Schulz, 2017). Participants

were excluded from analysis if they failed the control trials (n=29), parental interference (n=1), and fussiness (n=2), for not being English dominant (n=4), and for having cognitive or language impairments (n=2). This left 50 children included in the analysis (M=4.5, min=3.61, max=4.99)

Stimuli

The stimuli for the test trials were the same as the *small* stimuli of Experiment 1.

Procedure and Design

The experiment employed a two condition between-subject design. The *unprimed* condition was identical to the small numerosity condition of Experiment 1. The *primed* condition modified both training and test phases. After the four training trials, children were told they were going to look at families of animals and were asked to tell the experimenter what kind of family they saw. Children were then shown three scenes of four identical looking animals – four monkeys, four pandas, and four giraffes. After the children named the type of family for each scene, they were told they were going to look at additional pictures, and that they needed to tell the experimenter what kind of family they saw before the puppet described the pictures. In the test phase, the experimenter presented the visual stimuli and asked the child *"what kind of family is this?"* After the child responded, the experimenter said *"Great! Let's hear what [the puppet] says"* and played the pre-recorded test sentence which was the same as the sentences used in the unprimed condition and Experiment 1. The rest of the procedure was identical to Experiment 1.

Exclusion Criteria

We used the same criteria as Experiment 1. Twenty-nine children were subsequently excluded from analysis for failing the control trials.

Results

Since the key manipulation concerns how children will interpret *some-full* trials, and since children overwhelmingly accepted the *some-subset* trials (95% of the time), we focus our analysis on children's responses in the *some-full* trials, shown in Figure 4. The results show that children generated more implicatures (i.e, rejected the puppet's use of *some* in the *some-full* trials) more often in the *primed* condition than in the *unprimed* condition.

We analyzed these results using mixed effect logistic regression implemented in *R* using the *lme4* package. Our initial model estimated participant response (*accept* or *reject*) as a function of priming condition (*primed*, *unprimed*), age as a continuous variable, and the interaction between priming condition and age. The random effects structure included random intercepts for participants, with priming condition specified as random slopes against the participant intercepts.[6] We used backwards elimination to find the best model, which included only priming condition as a predictor. This model was significantly better than the null model ($\chi(1) = 16.66$, $p < .0001$). FDR-adjusted comparison tests showed that children in the primed condition were more likely to reject the puppet's statement than children in the unprimed condition ($\beta = -10.0$, $p < .0001$).

[6] Initial model formula: response ~ 1 + priming.condition + age.continuous + priming.condition:age.continuous + (1 + priming.condition |participant)

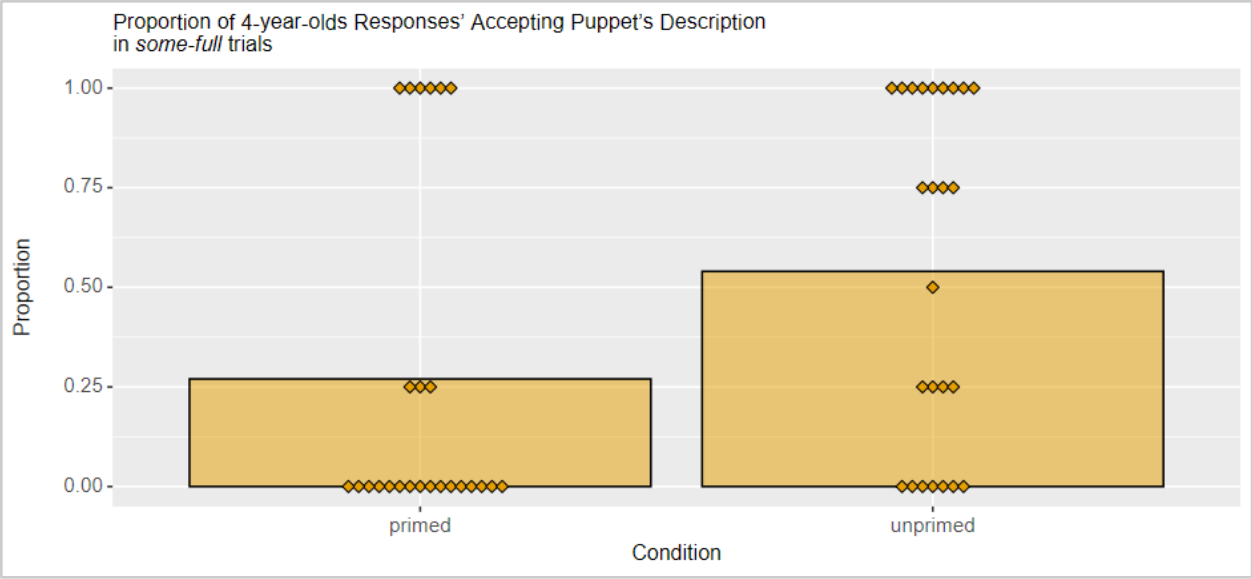

Figure 4: Mean proportion of 4-year-olds' responses that accepted the puppet's statement in some-full trials. The y-axis shows mean proportion, the x-axis the two between-subject priming conditions. Diamonds represent individual participants' mean response.

As in Experiment 1, given the relatively large number of children excluded from analysis for failing the control trials, we examined whether exclusion rates were even across the *primed* and *unprimed* condition, to ensure sure the greater rate of implicatures in the *primed* condition was not due to a greater exclusion rate. Exclusions were almost exactly equal: 14 out of 39 participants were excluded from the *primed* condition and 15 out of 40 were excluded from the *unprimed* condition.

Discussion

Our results are consistent with the prediction of the construal hypothesis, that in order to interpret *some* as meaning *some but not all*, the multiplicity needs to be construed as a group. 4-year-olds seem capable of doing this if they are given an explicit linguistic reason to view the multiplicity as a group, such as being told the group forms a *family.* Numerically, the rate of

implicature generation for 4-year-olds in the large numerosity condition in Experiment 1 and those in the family prime condition of Experiment 2 are similar (75% vs. 73%). The results of the present experiment suggest that the *large numerosity* condition in Experiment 1 facilitated implicatures because it promoted group construal rather than by merely making numerosity salient.

4. General Discussion

Experiments 1 and 2 investigated the role of group construal in preschoolers' ability to generate scalar implicatures with *some*. The construal hypothesis predicted that children will interpret *some* to mean *some but not all* when they are given reason to construe multiple entities as forming a group. In Experiment 1 we showed that large numerosities facilitated *not all* implicatures, with children who assessed twenty items generating more implicatures than those who assessed four items. This is consistent with the idea that large numerosities recruit ensemble representations, where small numerosities do not, and that ensemble representations facilitate group construal. In Experiment 2 we showed that in small numerosity contexts, a linguistic cue to construe the entities as a group like the collective noun like *family* helped children generate more *not all* implicatures. Experiment 2's results suggest that children's success in the large numerosity condition in Experiment 1 was not due to the large numerosity per se, but because those large numerosity contexts facilitated group construal, likely via ensemble representations. Together the two experiments show that group representations, independent of numerosity, is what matters for generating *not all* implicatures. Since our design did not contrast underinformative uses of *some* with *all*, children's implicatures were generated "out of the blue" – i.e., in the absence of any explicit contrastive information. Group construal seems sufficient for generating *not all* implicatures with *some.*

Our results suggest that in previous studies where children were found to generate *not all* implicatures, certain aspects of those studies' experimental design may have increased the saliency of the group relative interpretation of *some,* possibly by highlighting proportion. For example, in studies where children first assessed sentences with *all* before assessing sentences with *some* (Skordos & Papafragou, 2016; Foppolo et al., 2012; Pouscolous et al., 2007), having scenarios where *all* is false due to quantity may have highlighted proportion, rather than simply numerosity, since determining that sentences with *all* are false requires understanding that only a proper subset of entities have the described property. In studies where children had to select the scene or scenario best described by a sentence using *some* (e.g., Katsos & Bishop, 2011, Papafragou et al., 2018, Rees et al., 2023), the visual contrast between scenes may have highlighted differences in proportion. Since previous work did not seek to distinguish the relevance of numerosity versus proportion in the role of generating implicatures, we cannot be certain that proportion was being highlighted in those studies. However the results of our Experiment 2 suggest that proportion is what matters. Since proportions involve part-whole relations, they underlyingly involve representing multiple entities as forming a group.

Consequences for Implicature Development

An important implication of our results is that a key contributor to children's difficulty generating implicatures is their lack of understanding that when *some* is used in partitive syntactic frames, it unambiguously signals that the entities being must be construed as forming a group. In this they differ from the adults in our study, who generate implicatures even in the absence of other cues to group construal. Under the construal hypothesis, one thing that develops as children become more adult-like is their ability to use the morphosyntactic cues of their language to determine whether a phrase with *some* requires a group relative or multiplicity

asserting interpretation. That allows older children and adults to think of small numerosities as groups in the absence of non-morphosyntactic cues to do so. Previous proposals have generally assumed that children's literal interpretation of implicature triggering sentences is adult-like, and that components of the implicature computation mechanism are what is underdeveloped. For example, the ability to generate alternatives via lexical replacement (Barner et al., 2011; Gotzner, Barner & Crain, 2020) or the ability to determine what is the relevant dimension of sentence assessment (Skordos & Papafragou, 2016). Though generating the alternatives needed for implicature computation via lexical replacement may be one source of children's difficulty with implicatures, the results of Experiment 1 and Experiment 2 suggest that this cannot be the only source of difficulty. Due to the blocked design of our experiments, children did not hear the stronger alternatives using *all* when assessing uses of *some.* The manipulations that lead to children successfully generating implicatures also did not involve directly increasing the saliency or accessibility of the alternative propositions. If alternatives were made more salient, it was by virtue of increasing the saliency of the group relative interpretation. Similarly, while the ability to determine what is relevant in the conversational context is an important part of generating implicatures in an adult-like way, the results of Experiment 2 showed that children can succeed in situations where they aren't given contrasting information highlighting the relevance of quantity. The presence of cues to group construal seems sufficient to facilitate *not all* implicatures. Our results suggest that the basic components needed to generate implicature are in place by age 4, but that difficulty determining whether *some* should be interpreted as group relative or multiplicity asserting is one reason why they provide non-adult-like interpretations.

A broader consequence of this proposal is that implicature development should be understood as having important links to lexical and syntactic development. If lexical items are

semantically ambiguous in ways that influence the set of alternatives available for implicature generation, children must learn how to resolve these ambiguities in the way their target language requires in order to generate implicatures in an adult-like way. With respect to *some,* our results suggest that children need to learn the form-meaning mapping of partitive syntax to robustly generate implicatures with *some.* Children's non-adult understanding of the English partitive may be due to it being rare in their linguistic input. In plural contexts only around 5% of caregivers' uses of *some* use partitive syntax (Eiteljeorge et al., 2018). In addition only 6.5% of caregivers' uses of *some* more generally imply *not all* (Eiteljeorge et al., 2018)*,* meaning children are typically hearing *some* in contexts where it does not contrast with *all.* Non-linguistic factors, like a preference to attend to individuals in small numerosity contexts, or difficulty encoding proportional information in contexts with discrete entities, could also play a role.

If implicature development is linked to lexical and syntactic development, we expect different implicature-triggering lexical items to present their own particular word-learning challenges. As one example, modal expressions also generate scalar implicatures – *the penguin might be in the box* implies that it doesn't have to be in the box – and children are known to also have difficulty generating these implicatures (Novek, 2001; Ozturk & Papafragou, 2015; Shtulman & Philips, 2018). However, recent research has found that young children initially seem to treat modal words as marking distinctions of flavour instead of distinctions of force, i.e. they think *might* can be used as either a necessity modal or a possibility modal, but is restricted to epistemic contexts (Cournane et al., 2024). If *might* is treated as signalling either necessity or possibility, then sentences like *the penguin might be in the box* could be interpreted as literally true in situations when the penguin has to be in the box, and thus children would not generate the implicature that *might* implies *doesn't have to be*. Here again, children's non-adult-like

behaviour seems due to non-adult-like *semantic* knowledge rather than an inability to generate implicatures per se, suggesting that, more generally, we might expect non-uniform implicature development across different constructions.

Consequently, one important question for future work is to identify what potential challenges children face acquiring the semantics of other implicature-triggering lexical items like, *or, start, possible, warm,* and others. Cross-linguistic data on the morphosyntactic expressions of such implicature-generating constructions is one potential source of insight into these challenges, since they reveal the form-to-meaning mappings that any child has to be able to acquire. Data on the potential non-linguistic factors that influence children's interpretation of events, properties, possibilities, rules, and conjunctions are also likely to be informative about what interpretations children prefer entertaining.

Thinking and Talking about Quantities

In Experiment 1, we showed that a non-linguistic cue, increasing the overall set size of the entities being presented, facilitated the generation of *not all* implicatures in 4-year-olds.This finding aligns with recent research that has provided evidence of subtle behavioural differences in the interpretation of truth-conditionally equivalent quantifiers (Knowlton, Halberda, Pietroski & Lidz, 2023; Knowlton, Trueswell & Papafragou, 2023), and the more general idea that meanings can be thought of as instructions to build mental representations that interface with non-linguistic cognition (Pietroski et al., 2009; Knowlton, Trueswell & Papafragou, 2023). It is also consistent with recent suggestions that the alternatives used to generate implicatures must have a conceptual, and not simply logical, relationship to the utterance enriched by implicatures (Buccola et al., 2022; Smith, 2020). More generally, this result may indicate that linguistic

representations interface with other cognitive domains in ways that have an important consequence for how sentences are interpreted and pragmatically enriched.

If so, then the close link between group construal and the interpretation of *some* also suggest that language could provide a useful tool for investigating quantitative cognition. For example, in proportion comparison tasks, children and adults are known to attend to numerical information over proportional information in visual contexts that contain countable, discrete units (Boyer et al., 2008 , Jeong et al., 2007 ; Hurst & Cordes, 2018). It may be that for adults, linguistic factors like the use of partitive syntax make proportional information more salient than numerical information, by providing an important cue to construe the multiplicity as a group. Similarly, for children who have not learned the form-meaning mapping of partitive syntax, cues that help facilitate *not all* implicatures may also facilitate more accurate responses in proportional reasoning tasks. Future work can investigate the relationship between group construal and proportional reasoning, to better understand the links between pragmatic enrichment and quantitative cognition.

5. Conclusion

This paper investigated the role that the construal of the situation plays in the children's ability to generate scalar implicatures. Our experiments showed that children as young as 4 years old will interpret *some* to mean *some but not all* when given linguistic and non-linguistic cues to construe a multiplicity of items as forming a group. These results suggest that the basic components needed to generate implicatures are in place by age 4. As such, they point to the importance of investigating lexical, conceptual and morphosyntactic factors in addition to the pragmatic factors that might impede children's generation of implicatures.

6. References

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122–131.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*(2), 157–162.

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.

Benz, A., & Salfner, F. (2013). Discourse Structuring Questions and Scalar Implicatures. In G. Bezhanishvili, S. Löbner, V. Marra, & F. Richter (Eds.), *Logic, Language, and Computation* (Vol. 7758, pp. 35–50). Springer Berlin Heidelberg.

Boyer, T. W., Levine, S. C., & Huttenlocher, J. (2008). Development of proportional reasoning: Where young children go wrong. *Developmental Psychology*, *44*(5), 1478–1490.

Buccola, B., Križ, M., & Chemla, E. (2022). Conceptual alternatives. *Linguistics and Philosophy*, *45*(2), 265–291.

Carey, S. (2009). *The Origin of Concepts.* Oxford University Press.

Chierchia, G. (1998). Reference to Kinds across Language. *Natural Language Semantics*, *6*(4), 339–405.

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 *Experimental pragmatics* (pp. 283–300). Springer.

Choo, H., & Franconeri, S. L. (2014). Enumeration of small collections violates Weber's law. *Psychonomic Bulletin & Review*, *21*(1), 93–99.

Corblin F., Comorovski, I., Laca, B., & Beyssade, C. (2004). Generalized quantifiers, dynamic semantics, and French determiners. In Francis Corblin & Henriëtte de Swart (Eds.), *Handbook of French Semantics* (pp. 55–98). CSLI Stanford.

Cournane, A., Hirzel, M., & Hacquard, V. (2024). Mapping modal verbs to meanings: An elicited production study on "force" and "flavor" with young preschoolers. *Language Acquisition*, *31*(1), 57–80.

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

Davatz, J., & Stark, E. (2024). Romanian niște between non-specific and specific interpretations. In O. Kellert, S. Lauschus, & M. Rosemeyer, *Indefinites in Romance and beyond*. Language Science Press.

Degen, J. (2015). Investigating the distribution of *some* (but not *all*) implicatures using corpora and web-based methods. *Semantics and Pragmatics*, *8*, 11:1-55. https://doi.org/10.3765/sp.8.11

Degen, J., & Tanenhaus, M. K. (2015). Processing Scalar Implicature: A Constraint-Based Approach. *Cognitive Science*, *39*(4), 667–710.

Diesing, M. (1992). *Indefinites*. MIT Press.

Eiteljoerge, S. F. V., Pouscoulous, N., & Lieven, E. V. M. (2018). Some Pieces Are Missing: Implicature Production in Children. *Frontiers in Psychology*, *9*.

Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition*, *97*(3), 295–313.

Feigenson, L., Carey, S., & Hauser, M. (2002). The Representations Underlying Infants' Choice of More: Object Files Versus Analog Magnitudes. *Psychological Science*, *13*(2), 150–156. https://doi.org/10.1111/1467-9280.00427

Foppolo, F., Guasti, M. T., & Chierchia, G. (2012). Scalar Implicatures in Child Language: Give Children a Chance. *Language Learning and Development*, *8*(4), 365–394.

Foppolo, F., Mazzaggio, G., Panzeri, F., & Surian, L. (2021). Scalar and ad-hoc pragmatic inferences in children: Guess which one is easier. *Journal of Child Language*, *48*(2), 350–372.

Giusti, G. (2021). Partitivity in Italian. In G. Giusti & P. Sleeman (Eds.), *Partitive determiners, partitive pronouns and partitive case* (pp. 33–76). De Gruyter.

Gotzner, N., Barner, D., & Crain, S. (2020). Disjunction Triggers Exhaustivity Implicatures in 4- to 5-Year-Olds: Investigating the Role of Access to Alternatives. *Journal of Semantics*, *37*(2), 219–245.

Grice, H. P. (1967). Logic and Conversation. In P. Cole & J. L. Moran (Eds.), *Speech Acts* (Vol. 3, pp. 41–58). Academic Press.

Guasti, M. T., Chierchia, G., Crain, S., Foppolo, F., Gualmini, A., & Meroni, L. (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes*, *20*(5), 667–696.

Heim, I. (1983). File Change Semantics and the Familiarity Theory of Definiteness. In R.Bäuerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, Use, and Interpretation of Language* (pp. 164–189). De Gruyter.

De Hoop, H. (1995). On the Characterization of the Weak-Strong Distinction. In E. Bach, E. Jelinek, A. Kratzer, & B. H. Partee (Eds.), *Quantification in Natural Languages* (Vol. 54, pp. 421–450). Springer Netherlands.

Horn, L. (1972). On the Semantic Properties of Logical Operators in English [Ph.D., University of California, Los Angeles]. In *ProQuest Dissertations and Theses*.

Horowitz, A. C., Schneider, R. M., & Frank, M. C. (2018). The Trouble With Quantifiers: Exploring Children's Deficits in Scalar Implicature. *Child Development*, *89*(6), e572–e593.

Huang, Y. T., & Snedeker, J. (2009). Semantic meaning and pragmatic interpretation in 5-year-olds: Evidence from real-time spoken language comprehension. *Developmental Psychology*, *45*(6), 1723–1739.

Hurst, M. A., & Cordes, S. (2018). Attending to relations: Proportional reasoning in 3- to 6-year-old children. *Developmental Psychology*, *54*(3), 428–439. https://doi.org/10.1037/dev0000440

Hurst, M. A., & Levine, S. C. (2022). Children's understanding of most is dependent on context. *Cognition*, *225*, 105149.

Jeong, Y., Levine, S. C., & Huttenlocher, J. (2007). The Development of Proportional Reasoning: Effect of Continuous Versus Discrete Quantities. *Journal of Cognition and Development*, *8*(2), 237–256.

Katsos, N., & Bishop, D. V. M. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, *120*(1), 67–81.

Knowlton, T. (2021). *The Psycho-Logic of Universal Quantifiers*. [Ph.D., University of Maryland, College Park]

Knowlton, T. Z., Halberda, J., Pietroski, P., & Lidz, J. (2023). Individuals versus ensembles and "each" versus "every": Linguistic framing affects performance in a change detection task. *Glossa Psycholinguistics*, *2*(1).

Knowlton, T., Trueswell, J., & Papafragou, A. (2023). Keeping quantifier meaning in mind: Connecting semantics, cognition, and pragmatics. Cognitive Psychology, 144, 101584.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. MIT press.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*(2), 395–438.

Martí, L. (2008). The semantics of plural indefinite noun phrases in Spanish and Portuguese. *Natural Language Semantics*, *16*(1), 1–37.

Maule, J., Witzel, C., & Franklin, A. (2014). Getting the gist of multiple hues: Metric and categorical effects on ensemble perception of hue. *JOSA A*, *31*(4), A93–A102.

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

Ozturk, O., & Papafragou, A. (2015). The Acquisition of Epistemic Modality: From Semantic Meaning to Pragmatic Interpretation. *Language Learning and Development*, *11*(3), 191–214. https://doi.org/10.1080/15475441.2014.905169

Papafragou, A., Friedberg, C., & Cohen, M. L. (2018). The role of speaker knowledge in children's pragmatic inferences. *Child Development*, *89*(5), 1642–1656.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics–pragmatics interface. *Cognition*, *86*(3), 253–282.

Papafragou, A., & Tantalou, N. (2004). Children's Computation of Implicatures. *Language Acquisition*, *12*(1), 71–82.

Pietroski, P., Lidz, J., Hunter, T., & Halberda, J. (2009). The Meaning of 'Most': Semantics, Numerosity and Psychology. *Mind & Language*, *24*(5), 554–585.

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

Reed, A. M. (1991). On interpreting partitives. *In Bridges Between Psychology and Linguistics* (pp. 207–223). Psychology Press.

Rees, A., Carter, E., & Bott, L. (2023). Priming scalar and *ad hoc* enrichment in children. *Cognition*, *239*, 105572.

Selkirk, E. (1977). Some remarks on noun phrase structure. In P. W. Culicover, T. Wasow, & A. Adrian, *Formal syntax: Papers from the MSSB-UC Irvine conference on the formal syntax of natural language* (pp. 285–316). Academic Press.

Seržant, I. A. (2021). Typology of partitives. *Linguistics*, 59(4), 881–947.

Shtulman, A., & Phillips, J. (2018). Differentiating "could" from "should": Developmental changes in modal cognition. *Journal of Experimental Child Psychology*, *165*, 161–182.

Skordos, D., Feiman, R., Bale, A., & Barner, D. (2020). Do Children Interpret 'or' Conjunctively? *Journal of Semantics*, *37*(2), 247–267.

Skordos, D., & Papafragou, A. (2016). Children's derivation of scalar implicatures: Alternatives and relevance. *Cognition*, *153*, 6–18.

Smith, R. W. (2020). Similative plurality and the nature of alternatives. *Semantics and Pragmatics*, *13*, 15:1-44.

Sullivan, J., Bale, A., & Barner, D. (2018). Most Preschoolers Don't Know Most. *Language Learning and Development*, *14*(4), 320–338.

Wang, S. (2023). Effects of Processing Limits on Computing Scalar Implicatures: Evidence from Child English and Child Mandarin. *Language Learning and Development*, *19*(3), 345–368.