Non-adult Behavior of Children's Quantification in Logical Deduction Outside of the Language Domain

Rogério A. Guimarães Jr.
Department of Linguistics and Philosophy
Massachusetts Institute of Technology
Cambridge, MA, USA
rjunior@mit.edu

Abstract

Children in ages 4 to 6 present the exotic behavior of showing non-adult judgements of quantifiers, being specifically prone to *exhaustive pairing* errors. In the modern literature, the most accepted view about what factors cause this behavior suggest that pragmatical difficulties in understanding the context and the relevant objects to the sentence they need to evaluate, rather then semantic differences in the child and adult meaning of the quantifiers, are responsible for these non-adult judgements. If the factors are not semantic, but pragmatic, then we should also see the same type of non-adult judgement when children are presented with tasks that require the same type of deductive reasoning but do not involve the use of quantifiers. We propose an experiment to evaluate whether this is the case, and discuss the possible conclusions that we can draw from its results.

1 Introduction

It has well known to pysocholinguists that children at a certain age have non-adult behavior in the evaluation of quantifiers in a specific scenario. When presented a scene with objects of two distinct types A and B, in which all objects of type A show a property P, some but not all of the objects of type B do not show property P, but the remaining objects of type B do show property P, children diverge from adults in the evaluation of quantifiers like *"all"* and *"every"* when asked if all objects A show property P. Moreover, children point to the objects of type B that show property P to explain why the sentence is false. This type of error is called *Exhaustive Pairing*.

In the first studies to show this finding, Inhelder and Piaget (1958; 1964) showed children (ages 4-6) figures similar to Figure 1, with some blue circles, some blue squares, and some red squares of variable sizes. When asked *"are all circles blue"* children consistently answered negatively, and pointed out the blue squares as the reason.

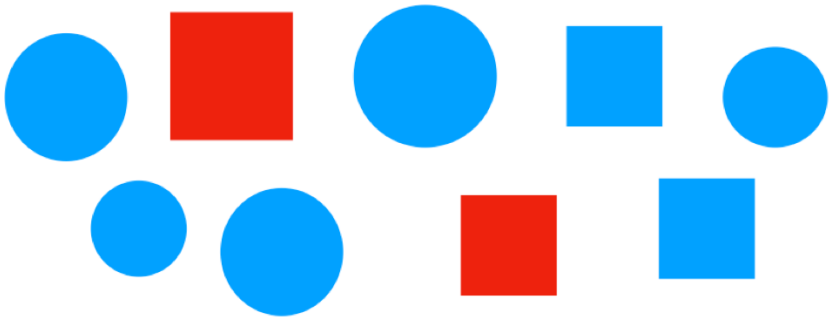

Figure 1: Example of scenario in which children frequently commit the error of Exhaustive Pairing.

When the researchers noted this behavior, their first assumption was that children had a non-adult semantic value for the quantifiers analyzed, in which symmetry is a necessary condition, i.e., *all circles are blue* is true if, and only if, all circles are blue and all blue entities are circle. However, this view

was disputed because such meaning for quantifiers would not be *conservative*, which is the only type of quantifier allowed in known natural languages (Barwise and Cooper, 1981). Furthermore, it has been shown that the behavior of children changes over time with respect to quantifiers (Aravind et al., 2017). Children become gradually more prone to commit over-exhaustive pairing over time between ages 4 and 7, but at the same time they become less prone to commit the inverse error of under-exhaustive pairing, which disproves the theory that they interpret the quantifiers as Inhelder and Piaget first proposed.

Other ideas have been proposed that justify this behavior at the semantic level in quantifiers. Philip (1995) proposed that children misinterpret *every* as a sentence-level quantifier that ranges over events rather than objects. However, it was later shown that children at the age of committing exhaustive pairing errors have a very good understanding of the core logical properties of *every* (Gualmini et al., 2003), which disputes the idea that quantifiers have a different meaning for children and adults.

If the problem is not in the semantics of the quantifiers, researchers proposed that it comes from pragmatic difficulties in understanding the relevant objects in the contexts presented, as if they were not able to restrict the universe of possibilities to the relevant objects and committed over-exhaustive errors for taking into consideration possible objects that are outside of the relevant context. Evidence for this view was provided by Drozd and van Loosbroek (2006) by showing that children perform much better in evaluating the quantifiers if they have some previous warm up story to familiarize them with the relevant context. Philip (2011) also provided support to this view by showing that the visual features of the scenes in the experiment, as well as the focus each object receives when a story is told, can be manipulated to nudge children towards the correct context interpretation, and thus also improve their performance in the quantifier evaluation task.

The idea that children have non-adult judgement of quantifiers because of pragmatic reasons related to context rather than semantic reason related to the meaning of the quantifiers has the implication that this non-adult judgement should not be restricted to the language domain. If children have a difficulty in understanding the relevant objects to a logical question rather than simply not understanding the words *all* and *every*, than we should not need to use these words in order to see this behavior. In principle, we should be able to come up with scenarios in which children commit exhaustive pairing errors even without the use of these quantifiers.

We propose an experiment in which children are compelled to evaluate the same logical statement represented by the sentence *"all circles are blue"*, but without the employment of the quantifier *all*. We predict that, in case exhaustive pairing errors are caused by pragmatic rather than semantic reasons, children should still commit the same mistake in our experiment.

In short, when evaluating the truth value of the sentence *"all circles are blue*, children are being asked to evaluate the logical statement described in (1).

$$\forall X \in \{x : x \text{ is a circle}\}, X \in \{y : y \text{ is blue}\} \tag{1}$$

We will present children with a story and with an assertion about this story that will be true if, and only if, statement (1) is true. If children maintain their non-adult responses, then the reasons that refrain than from correctly evaluating the *"all circles are blue"* sentence will not be restricted to the meaning of the quantifier, which would strengthen the view that such reasons are pragmatic.

2 Experiment

The experiment will consist of a story that will be told to children using illustrations. There are for variants of the story ending, and each variant will be told to a different group of children. Three of these variants will be control groups, to ensure the significance of the results, and the fourth variant will actually require the child to evaluate statement (1).

2.1 Participants

The experiment will be applied to 120 different children between ages 4 and 6. They will be divided into 4 groups with similar age distributions. Each group will be submitted to one of the experiments described in the following subsections.

2.2 Design

Individually, each children will be presented to the story described by Figures 2 to 4, and in each frame they will be told the following story:

- Frame 1: This is a story about some shapes that were having fun and dancing at a party! They were all very happy and the yellow start had a delicious apple.
- Frame 2: Suddenly the lights go off! Nobody knows who turned them off.
- Frame 3: When the lights come back, there is a surprise: someone stole the apple from the star! She got very upset and called the police to solve the theft. The policeman arrives and asks some questions to the star.

These frames and story lines will be told for the children in all four groups. After this each group will be shown a different ending of the story, to evaluate different phenomena.

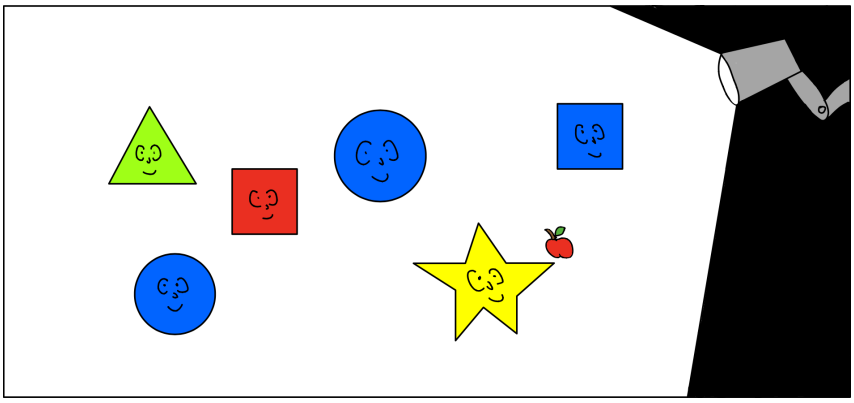

Figure 2: First frame of the story. Many shapes are having a party and star is holding an apple.

Figure 3: Second frame of the story. The lights of the party suddenly go off.

2.2.1 Presence of non-adult response

For the first group, the story will be over after these frames. After all the story is told, they will be simply asked *"are all circles blue?"*. This will be simply a control group to evaluate the presence of the non-adult response.

Since the previously mentioned studies showed that stories for familiarization of children with the scenes and context affect their ability to correctly evaluate quantifiers, it is important that the baseline be the percentage of children that still show non-adult responses even after they have been told the story.

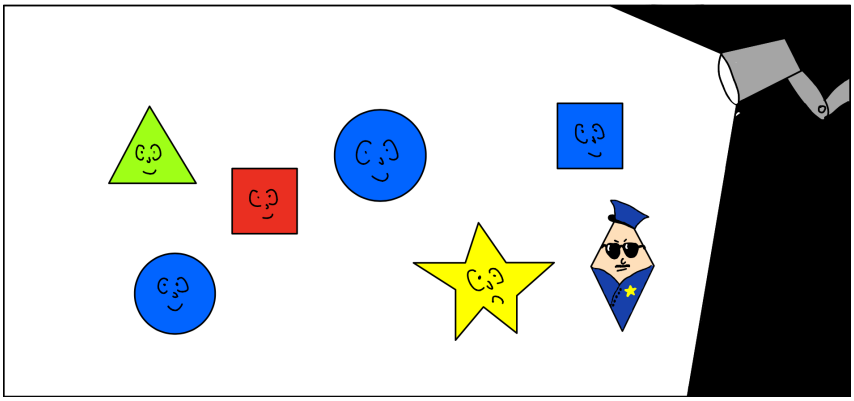

Figure 4: Third frame of the story. The lights come back on, but someone stole the star's apple. She calls the police to solve this theft.

2.2.2 Accuracy of logical deduction

For the second group, the story will continue and there will be a a description of a brief dialogue between the star and the policeman, as follows:

- **Policeman:** Did you see the color of the suspect?
- **Star:** No sir, it was too dark. But I could see the shape, it was a triangle!

After this dialogue, the children will be asked the color of the suspect. Since it was a triangle, and the only triangle at the scene is green, it is expected that they answer that the color of the suspect is green.

This experiment is a control group to certify that children have the logical ability to perform the deductive reasoning necessary to understand the situation that happened and deduct the color of the suspect from its shape, if it can be determined.

2.2.3 Engagement in logical deduction

For the third group, the story will continue and there will be a a description of a brief dialogue between the star and the policeman, as follows:

- **Policeman:** Did you see the color of the suspect?
- **Star:** No sir, it was too dark. But I could see the shape, it was a triangle!
- **Policeman:** Therefore, the suspect is green!

After this dialogue, the children will be asked whether the policeman is correct or not. Since the suspect is a triangle and the only triangle is green, the children are expected to answer that the policeman is correct.

This experiment is a control to certify that children are willing to engage in deductive reasoning to evaluate the answer of the policeman. Even though they might be able to perform the deduction, which will be assured by experiment 1, it is possible that they choose not to do it to evaluate the utterance of the policeman.

This experiment should guarantee us that children are indeed using their deductive reasoning to evaluate the sentence of the policeman, and therefore we can use their answers about this type of question to draw conclusions about their deductive reasoning.

2.2.4 Quantification Outside of the Language Domain

For the fourth group, the story will continue and there will be a a description of a brief dialogue between the star and the policeman, as follows:

- **Policeman:** Did you see the color of the suspect?

- **Star:** No sir, it was too dark. But I could see the shape, it was a circle!
- **Policeman:** Therefore, the suspect is blue!

After this dialogue, the children will be asked whether the policeman is correct or not. Since the suspect is a circle and all the circles in the context are blue, the children are expected to answer that the policeman is correct.

This is the experiment that will analyze the effect we are expecting. To recognize that the policeman is correct, the children must deduct that the policeman can only guarantee that the suspect is blue if all circles are blue. Therefore, they will have to evaluate the logical statement (1), even though quantifiers like *all* and *every* are never used in the story.

If the children answer that the policeman is correct, than they will be showing adult responses to logical statements similar to quantifiers in a situation outside of the language domain. If they answer that the policeman is incorrect, then they will be showing the same non-adult behavior when analyzing the same logical statement from a different domain, which is what we expect.

3 Discussion

There are four relevant possible outcomes for this study. We will discuss the implications of each one of them in the following subsections.

3.1 Failure in Control Experiments Scenario

The first three experiments are crucial controls to draw conclusions for the fourth experiment. In case children do note present the expected behavior in any of them, the results from Experiment 4 might become irrelevant.

The ways in which each control experiment can fail will be described in the following subsections.

3.1.1 Adult response in Experiment 1

In Experiment 1, children look at the third frame of the story and are asked *are all circles blue?*. This is intended to reproduce the results previously described from Inhelder and Piaget (1958; 1964), since we have the same situation in which all circles are blue, but we also have some blue and red squares. It is expected that a significant number of children present the non-adult behavior of answering *No* because of the presence of red squares.

Since the children will have some familiarization with the scene because of the story being told, it is possible that less children will present non-adult behavior then in other experiments without this familiarization. Still, if a significant number of students present the expected non-adult behavior, we will still have a baseline to compare against the outcomes of Experiment 4, which is the goal of Experiment 1.

3.1.2 Incorrect response in Experiment 2

If the children fail to answer that the suspect is green since it is a triangle, we will not be able to use experiment 4 to draw conclusions about their deductive reasoning because they will already be failing in a much more simple task.

We believe that Experiment 1 is simple enough that children will answer correctly. If this is not the case, we need further investigation. The researchers will ask children why they are giving their responses to understand why they are failing at such a simple task. With the information on what is misleading the children, we will have to redesign the study in order to use a task that they can actually solve.

3.1.3 Incorrect judgement in Experiment 3

If the children answer correctly in Experiment 1, but fail to recognize that the policeman is correct in Experiment 2, we also will not be able to draw conclusions about their deductive reasoning using Experiment 4. Even though the children might be able to correctly state the color of the suspect when asked, Experiment 4 is about the evaluation of a statement, and for any reason, children might treat theses tasks differently and simply not engage in deductive reasoning on them.

Therefore, it is necessary to guarantee that in a simple situation like Experiment 3, in which there is only one green triangle and thus green is trivially the color of the suspect, children do engage in deductive reasoning to evaluate the statement.

We decided to make both Experiments 2 and 3, instead of simply Experiment 3, in order to break down the possible sources of problems in case we have to redesign the experiment because children are failing in the control experiments. If children answer correctly in Experiment 2, but not in Experiment 3, we know that they are able to perform the necessary reasoning and we will only need to come up with a way to make them engage in such reasoning in the evaluation of statements.

3.2 Non-adult Response Scenario

In case children present a non-adult response in Experiment 4, this will be a strong evidence that pragmatical difficulties rather than semantic differences are the responsible for the non-adult interpretation of quantifiers in young children. This is the expected outcome of the study.

To asses whether children present non-adult responses in similar ways to when they are asked to evaluate quantifiers, we will compare the results of Experiment 4 to those of Experiment 1. Since Experiment 1 will be the baseline for the behavior of children with quantifiers, we will compare the frequency of children the present non-adult judgements in Experiments 1 and 4 and fit them into a chi-square distribution to define whether the two experiments have significantly different outcomes.

If the difference is not significant, it means that probably the same factors that cause children's mistakes in evaluation of quantifiers are also present in deductive reasoning outside of the language domain. Therefore, these factors cannot be related to the semantic value of the quantifiers, while pragmatic difficulties to understand context and relevant objects are a coherent explanation, since it should be present in both cases.

3.3 Adult Response Scenario

In case children present adult responses in experiment 4 at a statistically significant greater rate than in experiment 1, and answer correctly that the policeman is right in assert that the suspect is blue, this will present a challenge to our intuitions about the view that pragmatic difficulties are the responsible for the non-adult evaluation of quantifiers from young children.

There are two main explanations for such an outcome, which will be more deeply analyzed in the following subsections.

3.3.1 Non-adult Semantics of Quantifiers

The first explanation that can be raised for children presenting adult responses in Experiment 4, but not in experiment 1, is that the factors that create the non-adult response in experiment 1 are intrinsically related to the language domain, and specifically to the semantic value of the quantifier *all*.

This would raise all of the questions that were already presented to the previous authors that defend views relating exhaustive pairing errors to semantics.

However, we might not go as far as assuming that these factors must be semantic, but simply point out that they are restricted to the language domain. In short, we could conclude that in a task that evolves language, children are exposed to factors that produce non-adult responses but that are not present in deductive reasoning outside of the language domain.

If we assume that both tasks involve the evaluation of logic statement (1) by the children, we can only conclude that somehow this evaluation is being made differently in the two experiments. Maybe the brain performs logical evaluations in a different way or by a different module when they are related to quantifiers in natural language than when they are related evaluations of pure beliefs about the state of the world.

However, this is a very hard path to pursue. First, it is not easy to draw a hard line between what is *language domain* and what is *beliefs about the state of the world domain*. One of the primary functions of language is transmitting beliefs about the state of the world, and many of those most basic beliefs use quantifiers. I believe the best we can do in this path is explore the possibility that only attributing meaning to a quantifier is a logical evaluation that is performed by a separate module of the brain, or

in some other distinctive way, and that this specific system is somehow not fully developed in children. In fact, the existence of language universals in types of possible quantifiers might be a indicative that there is somehow a specific system in the brain that is responsible for them, and that this system might be separate from what evaluates other logical statements that we are able to produce, but that are never represented as quantifiers.

3.3.2 Mental Models and the Effect of Content

Another possible explanation for a significant divergence in the results of experiments 1 and 4 might be the way in which human beings actually perform deductions from logical statements.

Until now, we are simply considering that both tasks from experiment 1 and experiment 4 should be performed in the same way because they both correspond to a deduction, an inference from the same logical statement.

The view that we rely on a *Mental Logic* to reason asserts that in order to make a deductive reasoning, our minds recover the logical statement implied by the premises we have and use the rules of formal logic to draw conclusions from them (Braine and O'Brien, 1998). Our argument that Experiments 1 and 4 should have similar outcomes relies on the fact that they are both evaluating the same logical statement. To argue that the children will perform this evaluation in the same way just because it implies the same logical statement is somehow accepting the framework of mental logic, in which the same conclusions would be always drawn from formal logic if we have the same premises.

However, in the past few decades, this view has been challenged because of wrong predictions about the way humans make deductive inference (Johnson-Laird, 2010). In particular, humans make systematic mistakes in logical deduction that should not be the case if they were really using formal logic to draw their conclusions. Due to this, many psychologists support the view of *Mental Models* (Johnson-Laird, 2006). This view suggests that instead of using formal logic, people form possible iconic models that are possible given the premises and draw inferences from what is true or probably true in all of them.

If we assume Mental Models, it is possible that people draw different conclusions from the same logical statement when presented in different forms because the conclusions will have different probabilities of being true in the different mental models created for each situation. One example is the experiment performed by Oakhill (1989), which subjects were presented with two different scenarios. In the first scenario, they are given the following two pieces of information:

1. All of the Frenchmen in the restaurant are gourmets.

2. Some of the gourmets in the restaurant are wine-drinkers.

When the subjects were asked what, if anything, follows from these premises, most of the participants inferred that some of the Frenchmen in the restaurant are wine-drinkers.

However, another group of subjects was presented with the following two pieces of information:

1. All of the Frenchmen in the restaurant are gourmets.

2. Some of the gourmets in the restaurant are Italians.

When they were asked what, if anything, followed from these premises, only a very small portion of the participants replied that some of Frenchmen are Italian.

This happens because in the first scenario, it is much easier to come up with a hypothesis, a mental model in which the proposed conclusion is true. In the second scenario, the same conclusion is less likely because it is hard to imagine a mental model that encompasses it.

Therefore, it could be that, for some reason, children might answer experiment 4 more correctly because the situation to which they are presented makes it harder for them to come up with mental models that include the wrong conclusion. We could think, for example, that when asked whether all circles are blue, it is easy for children to come up with mental models in which there are other relevant objects, not presented in the context, that make the statement false (e.g. the existence of both blue and red squares

suggests that circles might also come in different colors, and there might also exist a red circle somewhere else). However, when they are trying to figure out who stole the apple from the star, it might be harder for them to come with mental models in which a non-blue entity stole the apple, given that they did not a possible candidate for that entity neither in the beginning nor in the ending of the party and they know that the theft took place somewhere in between.

4 Conclusion

We conclude that this experiment has a great potential of providing a better understanding of what factors affect children's ability to evaluate quantifiers. Specifically, if it reproduces their non-adult judgements in domains outside of language, it will be a very strong evidence that these factors really do not have to do with the semantic value of quantifiers, and most probably come from pragmatic difficulties. On the other hand, if the non-adult judgements are not reproduce, these results will provide a challenge to the view that these factors are pragmatic, but we have already provided possible ways to interpret such results without the need to assume that the factors are semantic. If this is the case, we would need further work to asses whether the differences in deductive reasoning inside and outside of the domain of quantifiers are being caused by differences in the mental models that they evoke in the subjects.

References

Athulya Aravind, Jill de Villiers, Peter A de Villiers, Christopher J Lonigan, Beth M Phillips, Jeanine Clancy, Susan H Landry, Paul R Swank, Michael Assel, Heather B Taylor, et al. 2017. Children's quantification with every over time. *Glossa*, 10(X):1.

Jon Barwise and Robin Cooper. 1981. Generalized quantifiers and natural language. In *Philosophy, language, and artificial intelligence*, pages 241–301. Springer.

Martin Braine and David P O'Brien. 1998. *Mental logic*. Psychology Press.

Kenneth F Drozd and Erik van Loosbroek. 2006. The effect of context on children's interpretations of universally quantified sentences. In *Semantics in acquisition*, pages 115–140. Springer.

Andrea Gualmini, Luisa Meroni, and Stephen Crain. 2003. An asymmetric universal in child language. In *Proceedings of Sinn und Bedeutung*, volume 7, pages 136–148.

Bärbel Inhelder and Jean Piaget. 1958. *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures*, volume 22. Psychology Press.

Birbel Inhelder and Jean Piaget. 1964. The early growth of logic in the child, trans. ea lunzer and d. papert.

Philip Nicholas Johnson-Laird. 2006. *How we reason*. Oxford University Press, USA.

Philip N Johnson-Laird. 2010. Mental models and human reasoning. *Proceedings of the National Academy of Sciences*, 107(43):18243–18250.

Jane Oakhill, PN Johnson-Laird, and Alan Garnham. 1989. Believability and syllogistic reasoning. *Cognition*, 31(2):117–140.

William Philip and Peter Coopmans. 1995. Symmetrical interpretation and scope ambiguity in the acquisition of universal quantification in dutch and english. *Ms., University of Utrecht*.

William Philip. 2011. Acquiring knowledge of universal quantification. In *Handbook of generative approaches to language acquisition*, pages 351–394. Springer.