



HAL
open science

Expectations of Processing Ease, Informativeness, and Accuracy Guide Toddlers' Processing of Novel Communicative Cues

Marie Aguirre, Mélanie Brun, Olivier Morin, Anne Reboul, Olivier Mascaró

► **To cite this version:**

Marie Aguirre, Mélanie Brun, Olivier Morin, Anne Reboul, Olivier Mascaró. Expectations of Processing Ease, Informativeness, and Accuracy Guide Toddlers' Processing of Novel Communicative Cues. *Cognitive Science*, 2023, 47 (11), 10.1111/cogs.13373 . hal-04288880

HAL Id: hal-04288880

<https://hal.science/hal-04288880>

Submitted on 16 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Running Head: Cognitive costs and benefits

Expectations of Processing Ease, Informativeness, and Accuracy

Guide Toddlers' Processing of Novel Communicative Cues

Marie Aguirre^{1*}, Mélanie Brun², Olivier Morin^{3,4}, Anne Reboul⁵, Olivier Mascaró²

¹ Faculty of Letters and Human Sciences, Institute of Language and Communication Sciences, University of Neuchâtel, Pierre-à-Mazel 7, Neuchâtel, 2000 Switzerland

² Université Paris Cité, CNRS, Integrative Neuroscience and Cognition Center, F-75006 Paris, France

³ Institut Jean Nicod, Département d'études cognitives, ENS, EHESS, CNRS, PSL University, UMR 8129

⁴ Max Planck Institute of Geoanthropology, Jena, Germany

⁵ Laboratory of Cognitive Psychology, UMR 7290, CNRS and Aix-Marseille University, Marseille, France

Author Note

The authors declare no conflict of interest.

Acknowledgements

We are grateful to all members of the Integrative Neuroscience and Cognition Centre for their help at all stages of this research, and to A. Couderc and P. Le Floch for their help with data collection. We also express our gratitude towards all of the families and daycare centers that participated in this project. This research was supported by two grants from the Agence Nationale pour la Recherche (PragTICAL, ANR-14-ACHN-0020; FoundTrust, ANR- 21-CE28-0017-01) and by a grant from the Fyssen foundation (IRIS: Infants' Representation of Informativeness in Social and non-social contexts) to Mascaró. This study contributes to

the IdEx Université de Paris ANR-18-IDEX-0001, and to the labex Empirical Foundations of Linguistics (ANR-10-LABX-0083).

* Corresponding author: Marie Aguirre, Faculty of Letters and Human Sciences, Institute of Language and Communication Sciences, University of Neuchâtel, Pierre-à-Mazel 7, Neuchâtel, 2000 Switzerland. Contact: marie.aguirre77@gmail.com

Abstract

Discovering the meaning of novel communicative cues is challenging and amounts to navigating an unbounded hypothesis space. Several theories posit that this problem can be simplified by relying on positive expectations about the cognitive utility of communicated information. These theories imply that learners should assume that novel communicative cues tend to have low processing costs and high cognitive benefits. We tested this hypothesis in three studies in which toddlers ($N = 90$) searched for a reward hidden in one of several containers. In all studies, an adult communicated the reward's location with an unfamiliar and ambiguous cue. We manipulated the processing costs (operationalized as inferential chain length) and cognitive benefits (operationalized as informativeness) of the possible interpretations of the cues. Toddlers' processing of novel communicative cues were guided by expectations of low processing costs (Study 1) and high cognitive benefits (Studies 2 and 3). More specifically, toddlers treated novel cues as if they were easy to process, informative, and accurate, even when provided with repeated evidence to the contrary. These results indicate that, from toddlerhood onward, expectations of cognitive utility shape the processing of novel communicative cues. These data also reveal that toddlers, who are in the process of learning the language and communicative conventions of people around them, exert a pressure favoring cognitive efficiency in communicative systems.

Keywords: pragmatics, communication, processing costs, informativeness, cognitive development, toddlerhood.

Expectations of Processing Ease, Informativeness, and Accuracy**Guide Toddlers' Processing of Novel Communicative Cues****1. Introduction**

Humans are expert communicators. Many animal species use signals to convey information, but humans' communicative practices stand out in their frequency, flexibility, as well as in the scope and nature of the information they convey (Csibra & Gergely, 2009; Harris, 2012; Sperber & Wilson, 1986; Tomasello, 2010). These properties are all the more remarkable given the ambiguity of human communicative means, such as gestures or words. Any communicative action or label could, in principle, carry an infinity of potential meanings (Quine, 1960). As a result, the interpretation of what people communicate amounts to searching through an unbounded hypothesis space. Yet, even infants and toddlers manage to use and interpret communicative cues whose meaning is highly variable, and context-dependent (e.g., Bohn et al., 2022; Falkum et al., 2017; Grosse et al., 2010; Grosse & Tomasello, 2012; Köder & Falkum, 2020; Mazzarella & Pouscoulous, 2021; Pouscoulous & Tomasello, 2020; Schulze et al., 2013; Schulze & Tomasello, 2015; Stephens & Matthews, 2014).

Many authors have argued that the process of solving the difficult problem of interpreting what others communicate can be guided by presumptions about the cognitive utility of the communicated information (e.g., Clark, 1996; Degen & Tanenhaus, 2019; Frank, 2016; Grice, 1975; Sperber & Wilson, 1986). Relevance theory provides a general characterization of such guidance. It posits that by providing information to their audiences, speakers imply that it is worth processing. As a result, the interpretation of what speakers intend to communicate is guided by a presumption of optimal relevance, defined in terms of cognitive (or contextual)

benefits, and processing costs. *Ceteris paribus*, the greater the cognitive benefits and the smaller the processing costs, the lower the cost/benefit ratio, and the greater the relevance (Sperber & Wilson, 2002). Here, we investigate how processing costs and cognitive benefits shape toddlers' processing of novel communicative cues. We assume that a cognitive benefit is a worthwhile difference to an individual's representation of the world – a true conclusion, for example (Wilson & Sperber, 2006). Throughout the paper, we focus on processing costs for addressees, namely on the efforts required from an addressee to process a communicative means. Drawing on Gricean and subsequent approaches (Frank, 2016; Grice, 1975; Sperber & Wilson, 1986), we hypothesized that when processing novel communicative cues, young learners assume that these cues have a low cost/benefit ratio (with high cognitive benefits and low processing costs).

1.1. Processing Costs

Sources of communicative messages attempt to modify the mental states of their audience. This goal is more efficiently achieved if the sources make the information that they intend to convey easy to produce (for the addresser), and easy to access and process (for the addressee). For instance, saying “it is two o'clock” is a more efficient way to convey that it is 2 PM than saying “the number of hours past noon is equal to the square root of four” because the former sentence is much easier to produce and process than the latter. As a result, two types of costs may shape the interpretation of communicated information: production costs (to the addresser), and processing costs (to the addressee). Several studies have investigated the role of addressers' production costs in shaping the interpretation of communicated information (e.g., Degen et al., 2013; Rohde et al., 2012). In contrast, there is little experimental data on the role of addressee's processing costs in shaping the interpretation of communicative cues. We addressed

this issue in Study 1, by investigating how processing costs shape learners' interpretation of novel communicative cues.

By offering information, sources of communicated messages imply that what they convey is worth being processed, and thus, that its cognitive benefits justify its processing costs (Sperber & Wilson, 1986). Indeed, adult speakers make efforts to reduce their audiences' processing costs. For example, individuals wearing digital watches that indicate the exact time are more likely to communicate the time using rounded, not exact, answers when this less precise information is relevant to their audience (Gibbs & Bryant, 2008; Van Der Henst et al., 2002).

Accordingly, processing costs are likely to play a central role in shaping the interpretation of communicated information. Several theories posit that the discovery of speakers' meanings can be guided by expectations about the ease of processing an utterance. Such expectations could originate from a maxim of manner which enjoins speakers to "be perspicuous" (Grice, 1975, p. 46), a presumption of relevance (Sperber & Wilson, 1986), a principle of least joint effort (Clark, 1996), or from speakers' tendency to strategically balance cost and informativeness (Bergen et al., 2016). Despite their differences, all these theories make converging predictions about the role of costs in shaping the interpretation of signals or utterances. They imply that, *ceteris paribus*, when people try to understand what is communicated to them, they are more likely to focus on hypotheses that are easier to process than on costlier ones.

We tested whether toddlers rely on comparable expectations by investigating whether they assume that novel communicative cues are easy to process (operationalizing ease of processing in terms of inferential chain length). Toddlers participated in a hiding game in which they had to discover the location of a reward. The experimenter provided information about the reward's location with an ambiguous novel communicative cue. We assessed whether toddlers

assumed that the novel cue indicated directly the location of the reward (and thus was easy to process) as opposed to indicating it indirectly (with unnecessary processing costs).

1.2.Cognitive Benefits

Benefits, and not just costs, shape learners' discovery of novel meanings. When speakers offer a piece of information to an audience, they imply that this information is worth processing, and thus, should yield some cognitive benefits. Indeed, many theories posit that expectations about cognitive benefits guide the interpretation of communicated information. These expectations are explicit in Grice's maxims of quality ("Try to make your contribution one that is true") and of quantity ("Make your contribution as informative as is required") (Grice, 1975, p. 45, 46). Several theories also postulate that expectations implying cognitive benefits guide the processing of communicated information (Degen & Tanenhaus, 2019; Frank, 2016; Frank & Goodman, 2014; Franke, 2011; Sperber & Wilson, 1986). In this study, we operationalized benefits in terms of informativeness, defined as the capacity of a stimulus to reduce uncertainty over a set of hypotheses (Frank & Goodman, 2012; Shannon, 1948).

Humans communicate in an informative manner from toddlerhood onward by adjusting their communicative behaviors to what their addressees already know (Liszkowski et al., 2008; Meng & Hashiya, 2014; O'Neill, 1996; O'Neill & Topolovec, 2001), and to what is expected or unexpected in a given context (Bannard et al., 2017). Moreover, even if toddlers and preschoolers often communicate ambiguously (Davies & Katsos, 2010; Nilsen & Graham, 2009), they can learn to avoid this, when prompted by clarification requests (Bahtiyar & Küntay, 2009; Matthews et al., 2007, 2012). Thus, by two to three years of age, children often adjust the informativeness of their linguistic and non-verbal communication to the needs of their

addressees. In short, from toddlerhood onward, children attempt to communicate in an appropriately informative manner.

On the reception side, preschoolers can evaluate the informativeness of utterances (Eskritt et al., 2008; Gweon et al., 2018; Gweon & Asaba, 2018; Morisseau et al., 2013). For instance, in Morisseau et al. (2013), 3- and 5-year-old children interacted with an adult who uttered instructions (e.g. “move the orange”). The context was manipulated so that instructions were either under-informative (e.g., there was more than one orange in view) or optimally informative (e.g., there was only one orange in view). By 3 years of age, children were more likely to gaze check their interlocutor when she was under-informative than when she was optimally informative. Moreover, by four years of age, children trust more and evaluate more positively informants that have been adequately informative rather than under-informative in the past (Eskritt et al., 2008; Gweon et al., 2018, experiment 4).

Preschoolers also rely on their sensitivity to informativeness when processing communicative cues such as labels (Bohn et al., 2022; Frank & Goodman, 2014), and when identifying the referent of an utterance (Stiller et al., 2015). For instance, in Frank and Goodman (2014), 3- and 4-year-old children were taught novel labels about objects features. An experimenter pointed at a target object with two features, while uttering a novel label (e.g., referring to “a dinosaur with a dax” while pointing at a dinosaur wearing both a headband and a bandana). Next to the object that was pointed at, there was a distractor object with only one of the two features (e.g., a second dinosaur wearing only a bandana). Due to the presence of the distractor, using the label “dax” to refer to the bandanna was not informative enough to pick out the dinosaur that was pointed at (since both dinosaurs wore a bandanna). Conversely, referring to the headband was an efficient way to target specifically the dinosaur that the speaker was

pointing at. By three years of age, children successfully inferred that the novel label referred to the unique feature, thus suggesting that they expected the speaker to be informative. Thus, by preschool age, expectations of cognitive benefits guide children's processing of familiar communicative cues (such as words and pointing). We built on this evidence to investigate how cognitive benefits may guide the processing of novel communicative cues during toddlerhood.

In short, we hypothesized that learners assume that the cost/benefit ratios of novel communicative cues are low, if not optimal, when processing them. We tested this hypothesis by assessing how toddlers' processing of novel cues is shaped by processing costs (Study 1) and benefits (Studies 2 and 3). Our Studies investigated the expectations guiding toddlers' processing of novel communicative cues. Importantly, we did not aim at testing whether these expectations are specifically triggered by communicative cues. In that respect, our Studies do not differ from most experiments on children's or adults' processing of communicated information (e.g., Behne et al., 2005; Bergelson & Swingley, 2012; Bohn et al., 2021; Frank & Goodman, 2012, 2014).

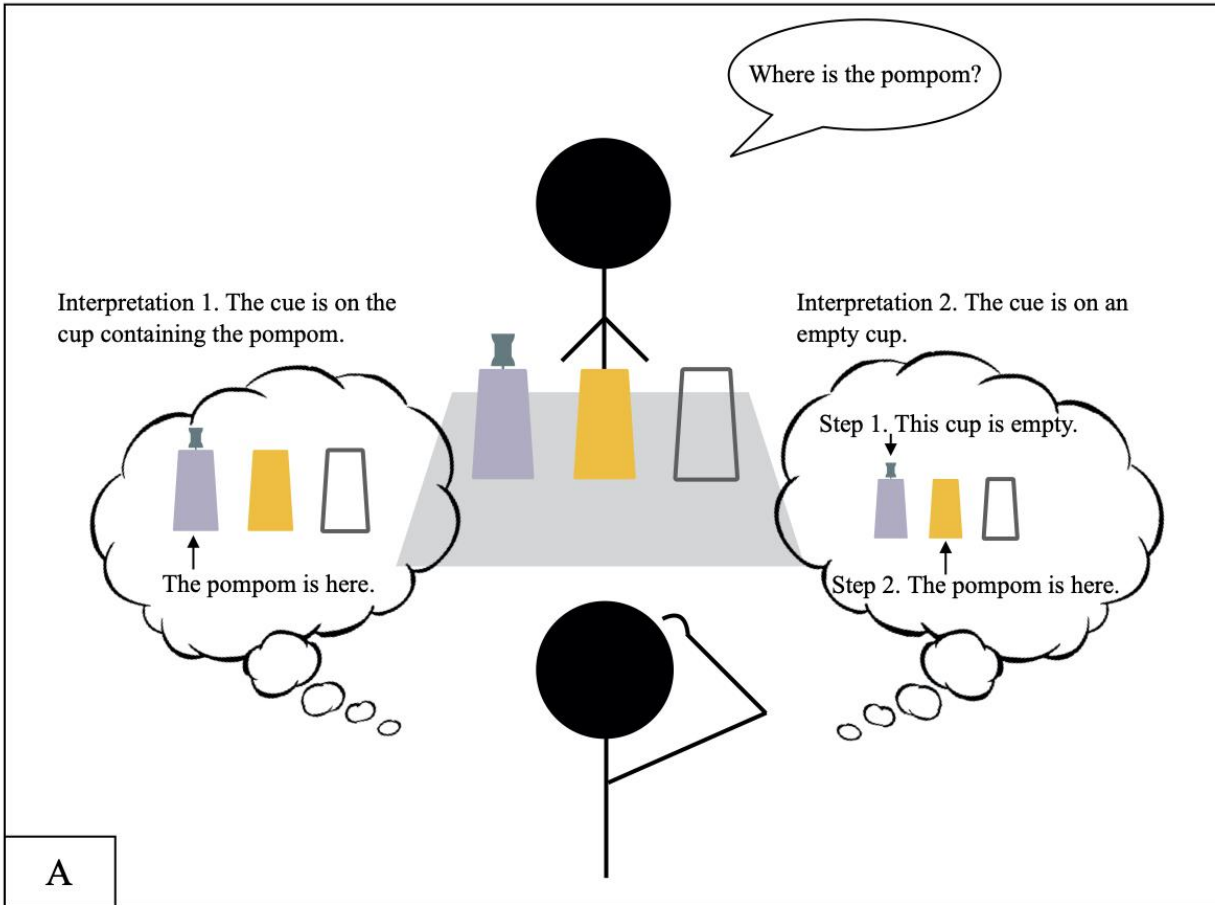
Study 1 tested whether toddlers rely on assumptions of low processing costs when processing novel communicative cues. In Study 1, participants were enrolled in a hiding game in which they had to discover the location of a reward hidden under one of three cups. One of the cups was transparent and could be excluded immediately; the other two were opaque. Before the participants could search for the reward, an experimenter gave them a clue about the reward location by placing a marker on top of one of the two opaque cups. In Study 1's ambiguous condition, the marker was deliberately ambiguous. It might communicate the reward location either directly (by being placed on the baited cup), or indirectly (by being placed on the empty cup that had to be excluded). The first way of processing the cue (marked cup = baited cup) is optimally easy to process (when learning that the reward is in cup A, one directly learns where

the reward is). The second way of processing the cue (marked cup = empty cup) is sufficient to discover the location of the reward (after excluding the marked cup, the only option left is the other opaque cup). However, it is costlier, because it requires the exclusion of one cup. If the reward can be in either cup A or B, upon learning that cup B is empty, one needs first to exclude cup B in order to discover the location of the reward (see Figure 1). We hypothesized that toddlers would resolve the marker's ambiguity by assuming that it was placed on the baited cup, thereby indicating the reward's location directly, at a low processing cost. Thus, we anticipated that toddlers would initially search for the reward under the marked cup, and might continue to do so throughout the experiment (although, in Study 1, the experimenter always placed the marker on the empty cup).

To control for participants' capacity to exclude the marked cup, we also tested a second group of toddlers on a negative condition. The latter condition was identical to the ambiguous condition, except that the experimenter explicitly uttered that the reward was not where she placed the marker. The aim of this control was to ascertain that participants were able to resist reaching for the marked cup (borrowing the logic of Palmquist et al., 2018). By 2.5 years of age—our target age group—the ability to use negative sentences to exclude a location is well-established (Austin et al., 2014; Feiman et al., 2017; Mascaro & Morin, 2015; Szabó & Kovács, 2022). Thus, we anticipated that participants would be less likely to reach for the marked cup in the negative condition than in the ambiguous condition. The data and analysis scripts of all studies are accessible on an open repository (https://osf.io/k9vwm/?view_only=d4c4ea1af6934e41840257a7625ca1ed).

Figure 1.

Two Possible Interpretations of the Marker in the Ambiguous Condition of Study 1.



Note: Panel A: Schematic representation of the stimuli in Study 2. The gray arrows indicate the locations of the pompom. The gray pin represents the marker placed by the experimenter. Panel B: Picture of actual experimental materials.

1.3.Ethical considerations

The research reported in this manuscript followed the guidelines of the Declaration of Helsinki. This study was approved by an independent committee for biomedical research (CPP Sud-Est II, IRB: 00009118) and by the local board of the daycare centers involved in the study. The parents of all participants gave their informed written consent prior to their inclusion in any of the studies (1–3).

2.Study 1

2.1.Method

2.1.1.Participants

Thirty-six two-year-old toddlers participated in Study 1, in a between-subject design (Ambiguous condition: $n = 18$, $M_{age} = 27.90$ months, $range = 20–35$ months, $SD = 4.67$, Negative condition: $n = 18$, $M_{age} = 28.8$ months, $range = 20–34$ months, $SD = 3.19$). Seven additional participants were excluded from the analysis either because they refused to complete the whole experiment (5), or because they failed to understand the task (i.e., they first chose an empty cup for two or more trials during the warm-up phase) (2). In Study 1 and in subsequent studies, our sample size was set to 18 participants per group. In previous studies in which 2.5 - year-olds had to discover the location of a hidden object using a marker, children discovered the object more often than predicted by chance, with effect sizes ranging from .902 to 1.998

(Tomasello et al., 1997; Zlatev et al., 2013). Assuming comparable effect sizes in our studies, a sample size of 18 was sufficient to reach a power ranging from .94 to 1 to compare the participants' choices of cups against chance with two-tailed tests (analyses performed with G-Power 3.1, Faul et al., 2007; with $\alpha = .05$).

For all Studies (1-3) the participants were native French speakers recruited from large French urban areas (> 1,000,000 inhabitants; for more details, see demographic information about the recruitment pool in the Supplementary Materials). We recruited participants by contacting a random sample of daycare centers that were selected to be within driving distance from our laboratory. For each daycare whose board validated the study, we invited families to participate in the project by sending an information letter to all the children falling within our target age ranges. We also occasionally tested participants in the laboratory to reach our target sample sizes. The participants tested in the laboratory were recruited by sending an invitation letter to a randomly selected sample of children born in our recruitment area. Each child participated in only one of the three studies reported in this paper. The participants were randomly assigned to conditions and studies. Data collection took place between May 2016 and June 2022 (with a break in 2019-2021), except for Study 1's negative condition (which was tested in 2023 following reviewers' comments). The participants were tested from Monday to Saturday, between 8:30 and 5:30 pm. Testing times were chosen with caregivers to ensure that the participants would not be too tired or sleepy during the study.

2.1.2. Materials

In all the studies, the participants were seated in front of the experimenter across a child-sized table, in a quiet room at their daycare center. A caregiver familiar to the participants was

present throughout the experiment. The caregiver was instructed not to influence the participants' choices at any time point during the experiment. In Study 1, the participants had to find a blue pompom (5 cm in diameter) hidden under one of the three upside-down plastic cups (Figure 1B). The three cups had the same shape and size (diameter: 6.5 cm, height: 11.5 cm); two cups were opaque (an orange one and a purple one); and the third cup was transparent. We introduced a transparent cup to make sure that participants were looking for the reward—as opposed to randomly selecting cups without caring about the reward's location. During the test trials, the transparent cup remained visibly empty; thus, we anticipated that participants would be less likely to choose the transparent cup than predicted by chance.

To communicate the location of the pompom, the experimenter used a “marker”—a plastic disc with a little protruding handle (the disc was 5 cm in diameter and 1 cm tall, the handle was a cylinder about 3 cm tall). This type of novel communicative cue has been validated in studies showing that (i) children rely on markers to locate objects from 18 months of age and (ii) that children treat markers differently from familiar communicative cues (Couillard & Woodward, 1999; Jaswal et al., 2010; Mascaro & Kovács, 2022; Mascaro & Sperber, 2019; Tomasello et al., 1997; Zlatev et al., 2013). During each warm-up and test trial, the cups were placed on a cardboard tray (64 cm × 27 cm) positioned on the child-sized table. The experimenter also used a large opaque screen made of white cardboard (45 cm x 48 cm x 9.5 cm) to occlude herself and the cups from the participants when she was hiding the pompom. A camera recorded the participants' behaviors.

2.1.3.Procedure

Warm-up Phase. At the beginning of the session, the experimenter showed three cups and a pompom to the participants. She explained that they would play a game in which they had

to find out where the pompom was hidden. Next, the participants were enrolled in three warm-up trials. At the beginning of each warm-up trial, the experimenter placed a rectangular cardboard tray on the table in front of her. She arranged the three cups in a row on the tray, approximately 15 cm apart. Next, she placed the pompom under one of the cups in full view of the child. She asked, “*Where is the pompom?*” before pushing the tray toward the participant to encourage them to search for the reward. The experimenter let the participant search for the pompom and congratulated them once they found it. Three consecutive warm-up trials were conducted. During each of them, the pompom was placed under a different cup and at a different location (to the left of the child, in front, or to the right). We assumed that the participants who chose an empty cup during two warm-up trials or more did not understand the task, and were, therefore, excluded from the data analysis. Following the warm-up phase, participants were enrolled in nine consecutive test trials.

Test Phase. Participants did not see where the reward was hidden at the beginning of test trials. Instead, the experimenter conveyed information about the location of the reward by placing a marker on top of the empty opaque cup. Thus, the test trials were identical to the warm-up trials, with the following exceptions. First, the experimenter did not show the participants where she hid the pompom. Instead, before hiding the pompom, the experimenter placed the large cardboard screen on the table between the cups and the participant. The screen was large enough to prevent the participant from seeing where the pompom was placed. The experimenter always hid the pompom under one of the two opaque cups.

Second, after hiding the pompom, the experimenter removed the screen from the test table, and asked, “*Where is the pompom?*” (Figure 1). Next, she held up the marker, and made eye contact with the participant, while saying, “*Look*”. Then, in the ambiguous condition, the

experimenter simply placed the marker on top of the empty opaque cup, without further verbal comment. In the negative condition, the experimenter manipulated the marker just like in the ambiguous condition, except that she said “*it’s not there*” while placing the marker on top of the empty opaque cup. In all conditions, the experimenter pushed the tray toward the participant and allowed them to search for the pompom. Participants’ first choice of a cup during each test trial was our primary dependent measure.

When the participant did not select the baited cup first, they were given a second opportunity to find the pompom. These second choices were coded separately from the participant’s first choices. We collected data on participants’ second choice simply to check that even participants who select the marked cup as their first choice can exclude it once they discover that it is empty. Nine consecutive test trials were conducted in Study 1. The location of each cup, of the baited cup (i.e., the cup that contained the pompom), and of the cup on top of which the marker was placed (to the left of the child, in front, or to the right) were counterbalanced across trials for each participant. The color of the opaque cup under which the pompom was hidden (orange or purple) was randomized across trials. Overall, the entire session (warm-up and test phase) lasted approximately five minutes.

2.1.4.Coding

We coded the participant’s first choice (i.e., which cup they lifted first) for all warm-up and test trials. For each test trial, we also coded the participant’s second choice of cup (in case they did not find the pompom in the first attempt). For each of these measures, we coded which cup the participant lifted first above the table. When the participant found the pompom in the first attempt, we coded their second attempt as missing data. All the data were double-coded by the experimenter and by a second coder, who was unaware of the hypotheses of the study,

double-coded 50% of the videos. There was a high correlation between the two raters' evaluation of the number of trials in which the participant selected the marked cup during their first or second attempts (Spearman's $\rho = 1, p < .001$).

2.1.5. Analyses

Statistical analyses were performed in the same manner for all Studies (1–3). Significance tests were always two-tailed. Analyses were computed using R (v. 4.3.1; R Core Team, 2022) on RStudio (v. 2023.06.1; Posit Team, 2023) with the following packages: dplyr (v. 1.1.2; Wickham et al., 2023), lme4 (v. 1.1-34; Bates et al., 2015), lmerTest (v. 3.1-3; Kuznetsova et al., 2017) and rcompanion (v. 2.4.30; Mangiafico, 2022). For all studies, when reported, the Wilcoxon signed-rank effect size refers to the matched rank biserial correlation of the “rcompanion” package.

In all of the studies, we assessed the effect of our factors using generalized linear mixed models (GLMMs) with the selection of a specific container (e.g., the cup with the marker) as a binary repeated measure (binomial distribution, logit link). In all of our models, we included age, sex and trial number as fixed effects to be controlled for, and subject identity as a random intercept. We included sex in our models because this variable has small yet consistent effects on toddlers' communicative abilities (Eriksson et al., 2012; Fenson et al., 1994). The likelihood ratio test (LRT) was used to compare the overall fit of distinct models. In order to compare the participants' performance to what was predicted by chance, we used non-parametric tests, given the discontinuous nature of our data.

2.2. Results

2.2.1. Warm-up Phase

Preliminary analyses confirmed that participants systematically chose the baited cup during the warm-up phase, in both conditions (see Supplementary Materials, Additional Results and Analyses: Study 1).

2.2.2. Test Phase: First Choice of Cup

When analyzing participants' first choice of a cup during test trials, we assessed the effect of Condition with three pairs of models, each pair being run on the selection of one specific cup (either baited, marked or transparent) as a binary repeated measure. For each of these analyses, we fitted (1) a null model which included the fixed effects of Age (in months), Sex (female vs. male) and Trial number (1–9), and subject identity as a random intercept; and (2) a full model, which added the fixed effect of Condition (ambiguous vs. negative) to the null model. Next, we performed analyses on the proportion of test trials in which each participant had selected each of the cups (baited, marked, or transparent) during their first attempt, per Condition (ambiguous and negative). We compared the proportion of trials in which participants first chose a given cup to what was predicted by chance—1/3, since there were three cups to choose from (see Figure 2).

Baited Cup. We first ran models on participants' choice of the baited cup. The resulting full model indicated that during Study 1's test trials, participants were significantly more likely to select the baited cup first in the negative condition than in the ambiguous condition ($\beta = 2.67$, $SE = 0.74$, $z\text{-value} = 3.61$, $p < .001$; LRT comparing null vs. full models: $\chi^2(1) = 13.30$, $p < .001$). The full model run on participants' first choice of the baited cup revealed no other significant effect.

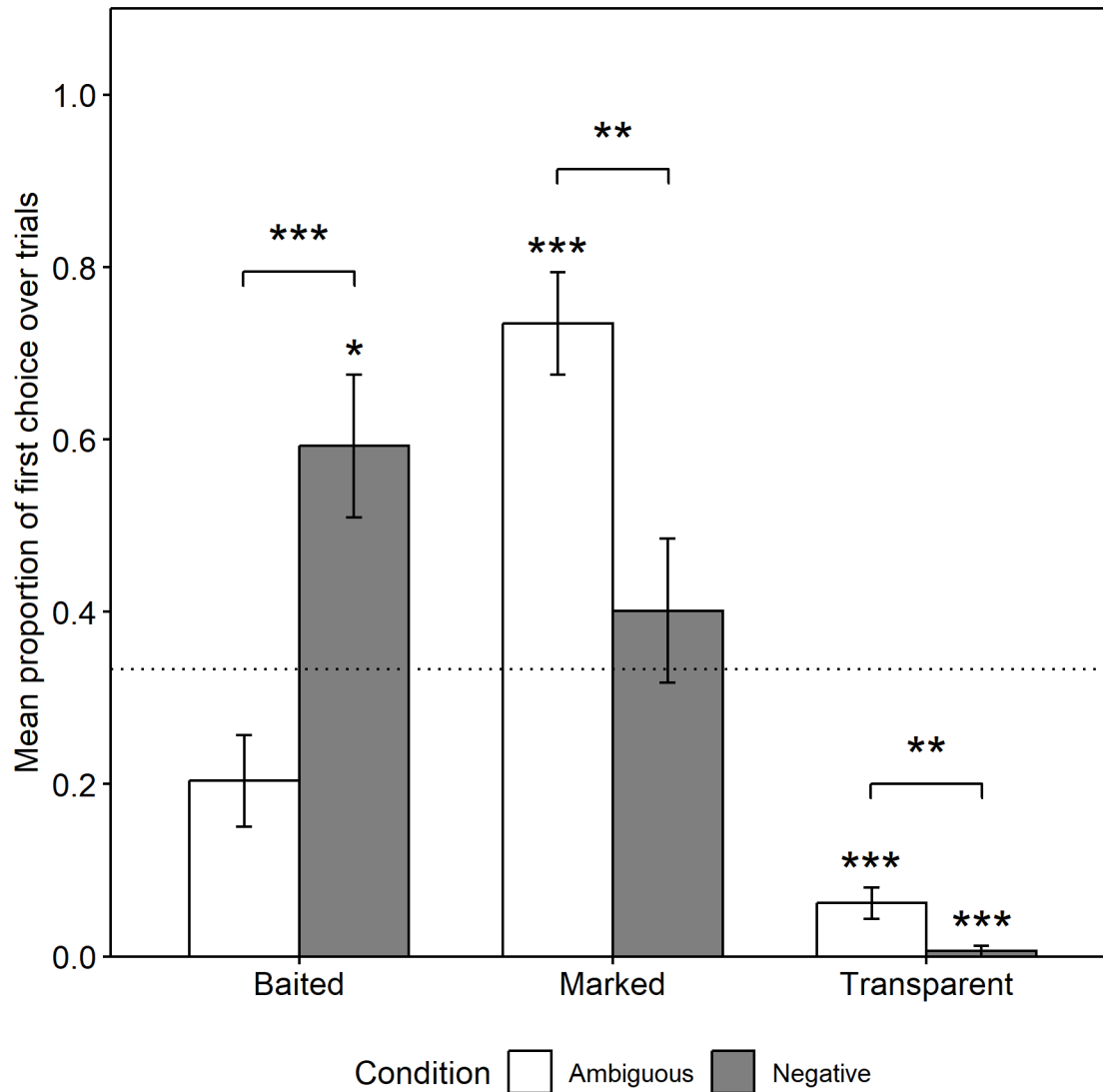
As shown in Figure 2, the average proportion of trials in which participants selected the baited cup first was significantly higher than predicted by chance (1/3) in the negative condition ($M = 0.59$, $M_{dn} = 0.67$, $SD = 0.35$, $W_+ = 143$, $p = .013$, $r = .67$, Wilcoxon signed-rank test). Conversely, the average proportion of trials in which participants selected the baited cup first tended to be lower than predicted by chance in the ambiguous condition ($M = .20$, $M_{dn} = 0.11$, $SD = 0.23$, $W_+ = 45$, $p = .08$, $r = -.47$).

Marked Cup. We also ran models on participants' choice of the marked cup. The resulting full model showed that Study 1's participants were significantly less likely to select the marked cup first in the negative condition than in the ambiguous condition ($\beta = -2.24$, $SE = 0.69$, $z\text{-value} = -3.23$, $p = .001$; LRT comparing null vs. full models: $\chi^2(1) = 9.98$, $p = .002$). The full model also revealed a significant effect of trial number showing that participants' tendency to select the marked cup first decreased significantly across trials in Study 1 ($\beta = -0.13$, $SE = 0.06$, $z\text{-value} = -2.26$, $p = .024$).

As shown in Figure 2, the mean proportion of trials in which Study 1's participants selected the marked cup first was significantly higher than that predicted by chance in the ambiguous condition but not in the negative condition (Ambiguous: $M = 0.74$, $M_{dn} = 0.78$, $SD = 0.25$, $W_+ = 164$, $p < .001$, $r = .92$; Negative: $M = 0.40$, $M_{dn} = 0.33$, $SD = 0.36$, $W_+ = 103$, $p = .46$, $r = .21$, Wilcoxon signed-rank tests).

Figure 2.

Performance in Study 1's Test Phase (First Choices).



Note: Average proportion of trials in which participants chose each of the cups (baited, marked, transparent) as their first choice, per condition (Ambiguous vs. Negative) during Study 1's tests trials. Error bars represent standard errors. The dotted lines represent the level of performance predicted by chance (i.e., 1/3). Comparisons of proportions against chance by one-sample

Wilcoxon signed-rank tests. Comparisons across conditions by GLMMs (detailed results and model descriptions are reported in the main text).

*: $p < .05$, **: $p < .01$, ***: $p < .001$.

Transparent Cup. The full model showed that Study 1's participants were significantly less likely to select the transparent cup as their first choice in the negative condition compared to the ambiguous condition ($\beta = -2.22$, $SE = 1.06$, $z\text{-value} = -2.09$, $p = .037$; LRT comparing null vs. full models: $\chi^2(1) = 7.26$, $p = .007$). The full model run on participants' first choice of the transparent cup revealed no other significant effect.

As shown in Figure 2, the proportion of trials in which participants selected the transparent cup first was significantly lower than predicted by chance in both conditions (Ambiguous: $M = 0.06$, $M_{dn} = 0$, $SD = 0.08$, $W_+ = 0$, $p < .001$, $r = -1$; Negative: $M = 0.01$, $M_{dn} = 0$, $SD = 0.03$, $W_+ = 0$, $p < .001$, $r = -1$, Wilcoxon signed-rank tests).

2.2.3. Test phase: Second Choice of Cup

In Study 1's test trials, when participants did not find the reward in their first attempt, they were given a second opportunity to search for it. During these second choices, the participants chose the baited cup more often than predicted by chance, in both conditions (see Supplementary Materials, Additional Results and Analyses: Study 1).

2.3. Discussion

During the test phase of Study 1's ambiguous condition, the novel cue used to communicate the location of the reward was ambiguous. It could be treated as indicating the location of the baited cup either directly (by marking it) or indirectly (by marking the empty opaque cup that needed to be excluded). Crucially, discovering the location of the reward by

excluding one empty location is costlier than learning directly where the reward is. Although the experimenter consistently placed the marker on the empty cup, toddlers tended to search for the reward under the marked cup. Thus, toddlers processed the marker as if it was easy to process. They assumed that the marker indicated the location of the reward directly, as opposed to indicating it indirectly, by signaling where the reward was not.

Children's tendency to select the marked cup in Study 1's ambiguous condition cannot be reduced to a low-level inability to resist reaching for the marked cup. In Study 1's negative condition, participants were explicitly told that the marker was placed where the reward was not. In that case, the participants still had to resist any putative drive to reach for the marked cup. Yet, toddlers were much less likely to choose the marked cup in the negative condition than in the ambiguous condition. Participants' second choices confirm that toddlers' tendency to choose the marked cup in Study 1's ambiguous condition cannot be reduced to an inability to resist reaching for the marked cup. Indeed, during their second choices, participants succeeded in locating the reward. Thus, when performing their second choices, Study 1's participants were able to locate the reward by excluding the marked cup after discovering that it was empty (in line with data from Gautam et al., 2021; Hill et al., 2012; Mody & Carey, 2016).

Similarly, the behaviors of Study 1's participants cannot be explained by a lack of interest in finding the reward. First, Study 1's participants were at ceiling in choosing the baited cup during the warm-up trials of both conditions. Second, during the test phase of Study 1, toddlers chose the transparent cup (which was visibly empty) much less often than predicted by chance in all conditions. Third, participants selected the baited cup more often than predicted by chance during the test phase of Study 1's negative condition. These results confirm that, in Study 1, participants cared about finding the reward, and did not chose cups randomly.

We doubt that when processing a novel cue, toddlers (or even adults) necessarily access and evaluate all of its possible interpretations. Instead, toddlers might consider just a small subset of a cue's possible interpretations. For instance, in Study 1's ambiguous condition, the experimenter implies that she will communicate information about the reward's location (by saying "*Where is the pompom? Look!*") before placing a marker on one opaque cup. When this communicative behavior is encountered for the first time, most toddlers probably consider only one hypothesis, namely, that the marker is placed on the baited cup. As evidence that this hypothesis is incorrect accumulates, some participant might start considering alternative ones. However, other participants might simply never revise their initial hypothesis. Thus, in Study 1's ambiguous condition, it is plausible that toddlers who kept selecting the cup with the marker never formed the hypothesis that the marker could be used to cue where the reward was not.

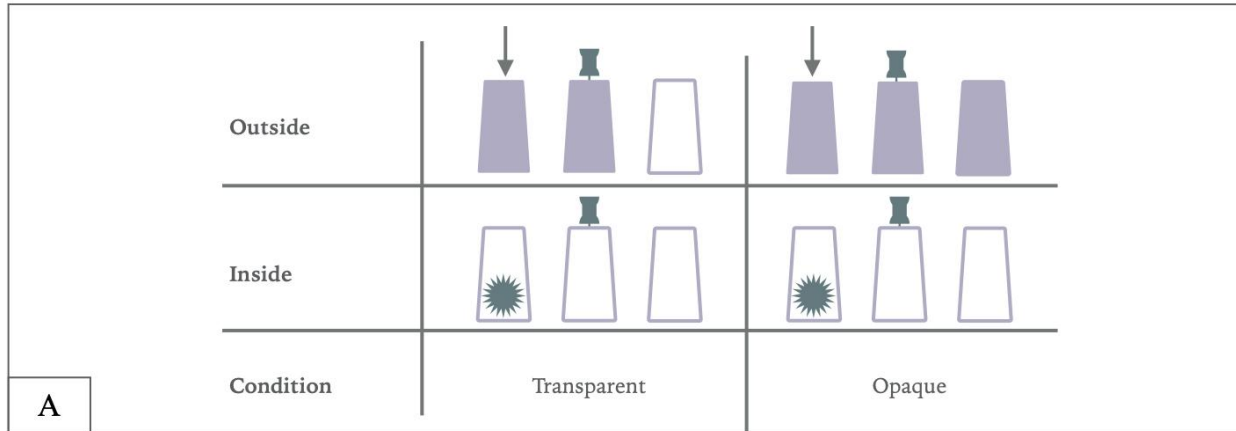
Study 1 suggests that toddlers' interpretation of novel communicative cues is shaped by expectations of low processing costs. Study 2 tested how cognitive benefits (the second component of cost/benefit ratios) shape toddlers' processing of novel communicative cues. In Study 2, cognitive benefits were operationalized in terms of informativeness. The participants were enrolled in a hiding game comparable to that of Study 1's ambiguous condition. However, we manipulated the context to modulate the informativeness of assuming that the marker was placed on the empty cup. We hypothesized that participants would be more likely to reach this interpretation when it was sufficiently informative to discover the location of the reward. Thus, our experimental design capitalizes on previous results showing that adults and preschoolers are more likely to find a negative sentence acceptable when its additional processing costs (compared to an affirmative one) are compensated by its cognitive benefits in terms of informativeness (Nordmeyer & Frank, 2018).

Study 2 was comparable to Study 1's ambiguous condition — children received information about the reward's location through a marker that was always placed on an empty cup. Yet, two main changes were introduced. First, we added a familiarization phase to reduce the participants' tendency to select the marked cup — since Study 1 evidenced an influence of cognitive costs that resulted in toddlers' tendency to select the marked cup. Second, in Study 2, we created two conditions in order to manipulate the informativeness of assuming that the cue marked the empty cup.

In Study 2's transparent condition, two of the cups used to hide the reward were opaque, while the third one was transparent (Figure 3A). Thus, in Study 2's transparent condition, assuming that the marker was placed on an empty cup was sufficiently informative to discover the location of the reward (by excluding the marked cup). In contrast, in the opaque condition, the three cups were opaque. In this case, assuming that the marker was placed on an empty cup was not sufficiently informative to discover the location of the reward (after excluding the marked cup, there were still two possible locations left for the reward). We assumed that if toddlers expect the cue to be informative, they should be more likely to assume that it marks the empty cup when this interpretation is sufficiently informative. As a result, we anticipated that the participants would be less likely to search for the reward under the marked cup in the transparent condition than in the opaque condition.

Figure 3.

Schematic Representation of Study 2's Stimuli Per Condition.



Note: Panel A: Schematic representation of Study 2's stimuli. The gray arrows indicate the locations of the rewards. The gray pin represents the marker placed by the experimenter. Panel B: Picture of actual experimental materials (transparent condition).

3. Study 2

3.1. Method

3.1.1. Participants

In Study 2, two-years-old toddlers were randomly assigned to one of two conditions (Transparent condition: $n = 18$, $M = 29.22$ months, $range = 23\text{--}35$ months, $SD = 2.92$; Opaque condition: $n = 18$, $M = 29.4$ months, $range = 23\text{--}37$ months, $SD = 3.79$). Nine additional participants were excluded from the analysis for the following reasons. Five participants refused to complete the experiment. Due to a recruitment error, in the opaque condition, 4 additional participants were tested beyond the planned 18 participants. We included data from these supernumerary participants in a complementary analysis (see the Supplementary Materials, Additional Results and Analyses: Study 2). This additional analysis showed that the pattern of significant results that we observed remained the same regardless of whether we included the supernumerary participants tested beyond the planned sample size.

3.1.2. Materials

The materials used in Study 2 were the same as those used in Study 1, with the following changes. All the opaque cups were of the same color (purple) and had a window (4 cm \times 4 cm) on their side, thus making it possible for the participants to see the contents of a cup when the window was facing them. In the transparent condition we used a fully transparent cup and two opaque ones, and in the opaque condition we used three opaque cups (Figure 3B).

3.1.3. Procedure

Warm-up Phase. The warm-up trials of Study 2 followed the same procedure as in Study 1 (after hiding the pompom in full view of the participant, the experimenter encouraged

them to search for it). During the warm-up phase, the cups' windows were turned away from the participant. The warm-up phase was followed by a familiarization phase.

Familiarization Phase. The familiarization trials of Study 2 followed the same procedure as the test trials of Study 1's ambiguous condition, except that the cups' windows were turned toward the participant, thus making it possible for them to see the cups' contents. Therefore, in each familiarization trial, the experimenter hid the pompom under one of the opaque cups, while an opaque cardboard screen prevented the participant from seeing where the pompom was placed. Next, the experimenter removed the screen from the test table and placed the marker on an empty opaque cup before encouraging the participant to search for the pompom. Crucially, during the familiarization trials, the participant could see the contents of the cups through the windows. Three consecutive familiarization trials were conducted, followed by the test phase.

Test Phase. The test trials of Study 2 followed the same procedure as Study 2's familiarization trials, except that the opaque cups' windows were turned away from the participant, thus making it impossible for them to see inside these cups. Nine consecutive test trials were conducted. Overall, the entire session (warm-up, familiarization and test phase) lasted approximately eight min.

During the warm-up, familiarization, and test phases of all the conditions, we counterbalanced the locations of the baited cup and of the cup on top of which the marker was placed (to the left of, in front of, or to the right of the child) across trials for each participant. During the warm-up, familiarization, and test phases of the transparent condition, we also counterbalanced the location of the fully transparent cup (to the left of, in front of, or to the right of the child) across trials for each participant.

3.1.4. Coding

We coded the participant's first choice (i.e., which cup they lifted first) for all warm-up, familiarization and test trials. For the test trials, we also coded the participant's second choice of cup (in case they did not find the pompom in the first attempt), just like in Study 1. All the data were coded by the first author and a second coder, who was unaware of the hypotheses of the study, double-coded 50% of the videos. There was a high correlation between the two raters' evaluation of the number of trials in which the participants selected the marked cup during their first and second choices (Spearman's $\rho = .98, p < .001$).

3.2. Results

3.2.1. Warm-up and Familiarization Phases

Participants chose the baited cup more often than predicted by chance during the warm-up and familiarization phases of Study 2, in both conditions (see Supplementary Materials, Additional Results and Analyses: Study 2).

3.2.2. Test Phase: First Choice of Cup

Just like in Study 1, we assessed the effect of Condition during test trials by running GLMMs on the participants' choice of a given cup as a binary repeated dependent measure. We fitted (1) null models, which included subject identity as a random intercept, and the fixed effect of Age (in months), Trial number (1–9) and Sex (female vs. male) and (2) full models, which added the fixed effect of Condition (transparent vs. opaque) to the null models.

Baited Cup. We first ran models on participants' choice of the baited cup. The resulting full model showed that Study 2's participants selected the baited cup first significantly more often in the transparent condition than in the opaque condition ($\beta = 3.11, SE = 0.66, z\text{-value} =$

4.74, $p < .001$; LRT comparing null vs. full models: $\chi^2(1) = 20.59$, $p < .001$). The full model also revealed a main effect of Age indicating older participants were more likely to select the baited cup first than younger ones ($\beta = 0.24$, $SE = 0.1$, $z\text{-value} = 2.5$, $p = .012$).

The average proportion of test trials in which Study 2's participants selected the baited cup first was higher than predicted by chance in the transparent condition, and lower than predicted by chance in the opaque condition (Figure 4; Transparent: $M = 0.57$, $M_{dn} = 0.56$, $SD = 0.37$, $W_+ = 136$, $p = .029$, $r = .59$; Opaque: $M = 0.15$, $M_{dn} = 0.11$, $SD = 0.16$, $W_+ = 13$, $p = .002$, $r = -.85$ Wilcoxon signed-rank tests).

Marked Cup. We also ran models on participants' choice of the marked cup. The resulting full model indicated that Study 2's participants were significantly less likely to select the marked cup first in the opaque condition than in the transparent condition ($\beta = -2.29$, $SE = 0.77$, $z\text{-value} = -3$, $p = .003$; LRT comparing null vs. full models: $\chi^2(1) = 9.10$, $p = .003$). The full model also revealed a main effect of Age indicating older participants were less likely to select the marked cup than younger ones ($\beta = -0.28$, $SE = 0.11$, $z\text{-value} = -2.49$, $p = .013$).

As Figure 4 shows, during Study 2's test phase, the average proportion of trials in which the participants first selected the marked cup was significantly higher than predicted by chance in the opaque condition ($M = 0.67$, $M_{dn} = 0.76$, $SD = 0.32$, $W_+ = 157$, $p = .002$, $r = .84$, Wilcoxon signed-rank test), but not in the transparent condition ($M = 0.40$, $M_{dn} = 0.33$, $SD = 0.37$, $W_+ = 95$, $p = .69$, $r = .11$, Wilcoxon signed-rank test).

In Study 2's transparent condition, the average proportion of trials in which the participants selected the marked cup first did not differ significantly from what was predicted by chance. Yet, children did not behave randomly in this condition. In fact, in the transparent condition, participants tended to use systematic strategies. One could consider that in Study 2's

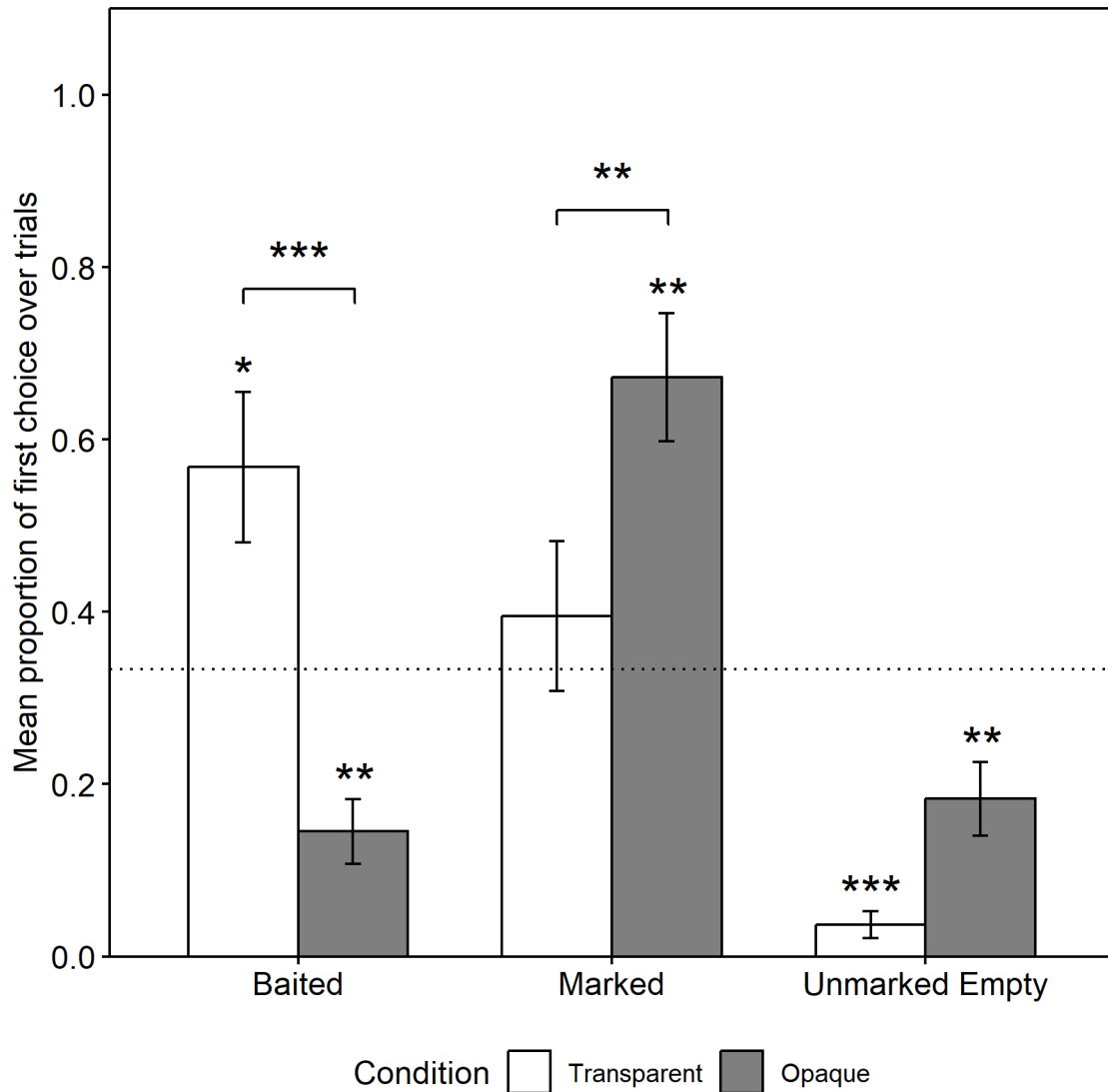
transparent condition, the participants used a systematic strategy if they selected one specific cup (e.g., the marked cup, or the unmarked opaque cup) more often than chance would predict, that is, for eight or more trials out of nine. More participants selected the marked cup systematically, in this sense, than what is predicted by chance (three participants out of 16, $p < .001$, $g = -0.31$, binomial test, with a proportion predicted by chance equal to .00097; for this analysis, we included only participants who participated in all trials, $n = 16$). Similarly, more participants selected the unmarked opaque cup systematically than what is predicted by chance (six participants out of 16, $p < .001$, $g = -0.13$, binomial test with a proportion predicted by chance equal to .00097). Finally, we assessed the proportion of participants who did not systematically avoid or select the marked cup (i.e., selected the marked cup in at least two trials, or seven trials at most). This proportion was significantly lower than that predicted by chance (seven participants out of 16, $p < .001$, $g = -0.06$, binomial test, a proportion predicted by chance equal to .86). Thus, in Study 2's transparent condition, children tended to use one of two strategies: either systematically selecting the marked cup or systematically selecting the unmarked opaque cup.

Unmarked Empty Cup. When analyzing Study 2's data we did not run GLMMs assessing the effect of condition on participants' choice of the unmarked empty cup because the unmarked empty cup differed across conditions (being transparent in the transparent condition, and opaque in the opaque condition). During the test phase of Study 2, the average proportion of trials in which participants selected the unmarked empty cup was higher than predicted by chance in both conditions (Figure 4; Transparent: $M = 0.04$, $M_{dn} = 0$, $SD = 0.07$, $W_+ = 0$, $p < .001$, $r = -1$; Opaque: $M = 0.18$, $M_{dn} = 0.12$, $SD = 0.18$, $W_+ = 23$, $p = .007$, $r = -.73$, Wilcoxon signed-rank tests).

In an exploratory analysis, we also compared performance in Studies 1 and 2. Study 2's transparent condition was modeled after Study 1's ambiguous condition, with the addition of a familiarization phase designed to increase participants' performance in finding the pompom. The exploratory analysis confirmed that this manipulation was effective (see Supplementary Materials, Additional Results and Analyses: Study 2).

Figure 4.

Performance in Study 2's Test Phase (First Choices)



Note: Average proportion of trials in which participants chose each of the cups (baited, marked, empty) as their first choice, per condition (transparent vs. opaque) during Study 2's tests trials. Error bars represent standard errors. The dotted lines represent the level of performance predicted by chance (i.e., 1/3). Comparisons of proportions against chance by one-sample Wilcoxon signed-rank tests. Comparisons across conditions by GLMMs (detailed results and model descriptions are reported in the main text).

*: $p < .05$, **: $p < .01$, ***: $p < .001$.

3.2.3. Test Phase: Second Choice of Cup

In Study 2, participants selected the baited cup as their second choice more often than predicted by chance in both conditions (see Supplementary Materials, Additional Results and Analyses: Study 2).

Discussion

Study 2's results indicate that toddlers' processing of unfamiliar communicative cues is guided by expectations of informativeness. In the test phase of Study 2's opaque condition, the participants kept searching first for the reward under the marked cup. Indeed, in Study 2's opaque condition, assuming that the marked cup was empty was not informative enough to discover the location of the reward. By contrast, in the transparent condition of Study 2, excluding the marked cup was sufficiently informative to discover the location of the reward. As predicted, the participants were more likely to exclude the marked cup in this condition than in the opaque condition. This result is all the more remarkable given that the marker was always used in the same manner—only the context differed across conditions.

Moreover, it is important to note that toddlers did not behave randomly in Study 2's transparent condition. They relied on one of two distinct strategies: either systematically excluding the marked cup, or systematically choosing it. This result indicates that a subset—but not all—of our participants reached the accurate interpretation of the marker in Study 2's transparent condition.

Study 2's data confirm that toddlers' tendency to rely on the marker in Studies 1-2 did not originate from an inability to exclude the marked cup. First, toddlers had no difficulties to disengage from the marker. In Studies 1-2, they performed at ceiling when excluding the marker on their second choice, once they had discovered that the marked cup was empty. Second, if

toddlers' tendency to select the marked cup had been unconditional, toddlers would have behaved similarly in the transparent and in the opaque conditions of Study 2.

Note that Study 2's primary conclusions are based on the effect of condition on participants' tendency to choose the marked cup first. In Study 2's test phase, condition also had an effect on participants' tendency to choose the baited cup first. Yet this effect was not as theoretically informative as the effect of condition on participants' choice of the marked cup. Understandably, after excluding the marked cup, participants could systematically choose the (unique) remaining opaque cup in the transparent condition. By contrast, after excluding the marked cup, participants had to choose randomly between two unmarked opaque cups in the opaque condition, thus explaining why they were less likely to select the baited cup in that condition than in the transparent condition.

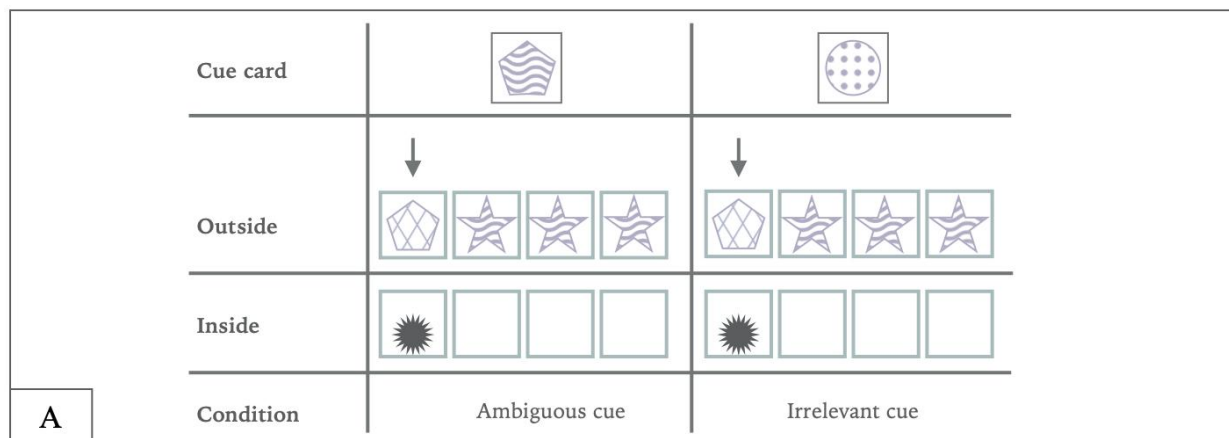
Studies 1-2 focused on "indexicals" (Burks, 1949), i.e., signs referring to some objects by means of a direct spatio-temporal connection. Study 3 investigated whether toddlers' expectations of informativeness can be found for novel communicative cues that differ from indexicals. In Study 3 we assessed whether expectations of informativeness guide toddlers' processing of visual icons. Study 3's ambiguous cue condition tested whether toddlers would process a novel ambiguous visual symbol as if it were informative. In Study 3's ambiguous cue condition, the participants had to find a reward hidden in one of four boxes on which pictures were glued (see Figure 5). The same picture was glued to three of the boxes (henceforth, the repeated picture), and a second picture was glued to the fourth box (henceforth, the unique picture). Before the participants could search for the reward, the experimenter showed them an ambiguous cue card. The cue card shared one feature with the unique picture and another feature with the repeated picture. For instance, the cue card might have the same shape as the unique

picture, and the same pattern as the repeated picture (we counterbalanced across trials which feature — pattern or shape — the cue card shared with the unique and repeated pictures). In short, the cue card was ambiguous, in that it could be taken to refer to the box with the unique symbol, or to the three boxes with the repeated symbol. Assuming that the cue card referred to the box with the unique symbol was more informative—it maximally reduced the participants' uncertainty about the reward's location. Conversely, assuming that the cue card referred to the boxes with the repeated symbol was less informative — since there were several boxes with a repeated symbol to choose from. We anticipated that toddlers would assume that the cue card was informative enough to discover the reward, and thus would search for the reward inside the box with the unique picture stuck to it.

Furthermore, to test whether our participants answered on the basis of the cue card in the ambiguous cue condition, we also used an irrelevant cue control condition. In the latter, the experimenter communicated with a cue card whose picture shared no common feature with any of the pictures on the boxes. Thus, in the irrelevant cue condition, the cue card carried no relevant information about the location of the reward. Accordingly, in this second condition, we anticipated that toddlers would be no more likely than chance would predict to search for the reward inside the box with the unique picture.

Figure 5.

Study 3's Stimuli.



Note: Panel A: Schematic representation of the stimuli in Study 3. The gray arrows indicate the locations of the pompom. Panel B: Picture of actual experimental materials.

4. Study 3

4.1. Method

4.1.1. Participants

In Study 3, two-years-old toddlers were randomly assigned to one of two conditions (Ambiguous cue condition: $n = 18$; $M_{age} = 28.94$ months, $range = 24-37$ months, $SD = 3.83$; Irrelevant cue condition: $n = 18$, $M_{age} = 28.83$ months, $range = 24-36$ months, $SD = 4.00$). Ten additional participants were excluded from the analysis for the following reasons: refusal to complete the whole experiment (5), failing to understand the task, that is, choosing an empty box for two or more trials during the warm-up phase (3), and technical issues (2).

4.1.2. Materials and set-up

The materials used in Study 3 were the same as those used in Study 1, with the following changes. Instead of searching for the reward inside the cups, the participants had to find the pompom in one out of four upside-down wooden boxes of identical size and shape (8 cm \times 8 cm \times 10 cm each). In Study 3 we used 4 containers (unlike in Studies 1-2) in order to increase the difference of informativeness between the two possible interpretations of the cue in the ambiguous cue condition. During the warm-up phase, plain boxes without pictures were used. During the test phase, each box had a picture glued on its front face, and the experimenter communicated with the child using cue cards (7 cm \times 8 cm) on which pictures had been drawn. Each picture had a shape and was filled with a colored pattern. To create all the pictures used in the experiment, we used eight different shapes (circle, crescent, parallelogram, pentagon,

rhombus, rounded-corner triangle, square, and star) and eight different patterns (small confetti, colored dots, diamonds, Scottish tartan, waves, interweaved patterns, white diagonal bricks, white dots on a dark background). Young children tend to rely more heavily on shape than on color information in several learning contexts, a phenomenon sometimes called “the shape bias” (e.g., Diesendruck & Bloom, 2003; Landau et al., 1988). In Study 3, we aimed to counteract the shape bias by contrasting patterns (rather than plain colors) with shapes. Patterns are composed of small shapes; thus, they are more likely than mere plain colors to have equal salience to external shapes.

4.1.3.Procedure

Warm-up Phase. The warm-up phase procedure was the same for all conditions and it was identical to the warm-up phase of Study 1, with the following exceptions. First, there were four warm-up trials (instead of three). Second, the experimenter did not hide the reward under the cups, but under one of four identical plain boxes (without pictures on their front side). The location of the box containing the pompom on the cardboard tray (outer left, central left, central right, or outer right position) was counterbalanced across trials within participants. The warm-up phase was followed by a test phase.

Test Phase. Eight consecutive test trials were divided into two blocks of four trials. The test trials were identical to the warm-up trials, with the following exceptions.

First, during the test phase, pictures were glued to the front face of the boxes used to hide the reward. The front faces of the boxes were turned toward the participant throughout the test phase. For each test trial, one picture—repeated picture—was glued to three of the boxes. A different picture—the unique picture—was glued to the fourth box (see Figure 5). The shape and pattern of the unique and repeated pictures varied from trial to trial. For each participant, we used

four different unique pictures and four different repeated pictures in each of the two blocks of test trials. Each picture glued on the boxes was used as a unique picture for half of the participants, and as a repeated picture for the other half. In each test trial, all the pictures glued on the boxes were of the same color (green, blue, purple, or orange).

Second, during Study 3's test phase, the experimenter did not show the participant where she hid the pompom. Instead, before hiding the pompom, the experimenter placed the large cardboard screen on the table between the cups and the participant. In all conditions, the experimenter always hid the pompom under the box with the unique picture. We randomized across trials the location of the box with the unique picture on the cardboard tray (outer left, central left, central right, or outer right position). After hiding the pompom, the experimenter removed the screen from the test table and asked, "*Where is the pompom?*" Next, she communicated with the child by holding up a cue card, waited for an eye contact with the participant, looked at the participant in the eye while saying, "*Hey, look!*" while pointing at the picture on the card.

The pairings between the cue cards and boxes differed across conditions. In the test trials of the ambiguous cue condition, the picture on the cue card was ambiguous because it shared a different common feature with the unique picture and with the repeated pictures on the boxes. For each participant, in half of the trials, the picture on the cue card had the same shape as the unique picture, and the same filling pattern as the repeated pictures; in the other half of the trials, the opposite was true. Thus, in the ambiguous cue condition, the picture on the cue card could be used to refer either to the box with a unique picture, or to the three boxes with repeated pictures. However, the most informative interpretation of the cue card was that it referred to the unique

picture, which resulted in a maximal reduction of the participants' uncertainty about the location of the reward (Figure 5).

In the test trials of the irrelevant cue condition, the picture on the cue card had no common feature with any of the pictures glued on the boxes, that is, its shape and filling pattern were different from those of the unique and repeated pictures. For instance, in the example shown on Figure 5A, a circle with dots is drawn on the irrelevant cue card, while on the boxes there are either stars with waves (repeated pictures), or a pentagon with lines (unique picture). Thus, in the irrelevant cue condition, the cue card provided no information about the location of the reward. To achieve this result, we used the same sets of boxes and cue cards as in the ambiguous cue condition, and simply modified which cue card was paired with each given set of boxes.

In all conditions, after showing the cue card to the participant, the experimenter pushed the cardboard tray toward them, while asking again, “*Where is the pompom?*” Next, the participant was left to freely search inside the boxes until they found the pompom. At that point the trial ended. In case the participant chose the wrong box, the experimenter asked again, “*Where is the pompom?*” and added “*Look!*” while pointing at the picture on the card. Overall, the entire session (warm-up and test phase) lasted approximately fifteen minutes.

4.1.4.Coding

For each trial, we coded which box the participants chose first, that is, which box they lifted first above the table. In a few trials, Study 3's participants lifted two containers almost at the same time (5 trials out of 432). In these cases, we used frame-by-frame offline coding from video recordings to determine which of the two containers the participant had lifted first. The participants received a score of 1 for each trial in which they selected the box with a unique

picture and 0 otherwise. All the data were coded by the first author, and a second coder, who was unaware of the hypotheses of the study, double-coded 50% of the videos. There was a high correlation between the two raters' evaluation of the number of trials in which the participants had selected the box with the unique picture (Spearman's $\rho = .95$, $p < .001$). Disagreements were resolved through discussion.

4.2.Results

4.2.1. Warm-up Phase

Participants chose the baited container more often than predicted by chance during the warm-up phase of Study 3, in both conditions (see additional results and analyses: Study 3 in the Supplementary Materials).

4.2.2. Test Phase

To assess the effect of Condition we ran GLMMs with the participants' choice of the container with the unique picture as a binary repeated dependent measure. We fitted (1) a null model, which included subject identity as a random intercept, and the fixed effects of Age (in months), Trial number (1–9), Sex (female vs. male); (2) a full model, which added the fixed effect of Condition (Ambiguous vs. Irrelevant cue) to the null model. Comparing the null model to the full model revealed that adding Condition to the null model improved its goodness of fit by a LRT ($\chi^2(1) = 3.94$, $p = .047$). The full model revealed a significant main effect of Condition ($\beta = -0.52$, $SE = 0.27$, $z\text{-value} = -1.97$, $p = .049$) showing that children tended to select the box with the unique picture more often in the ambiguous cue condition than in the irrelevant cue condition. In addition, the full model revealed a main effect of Age indicating older participants

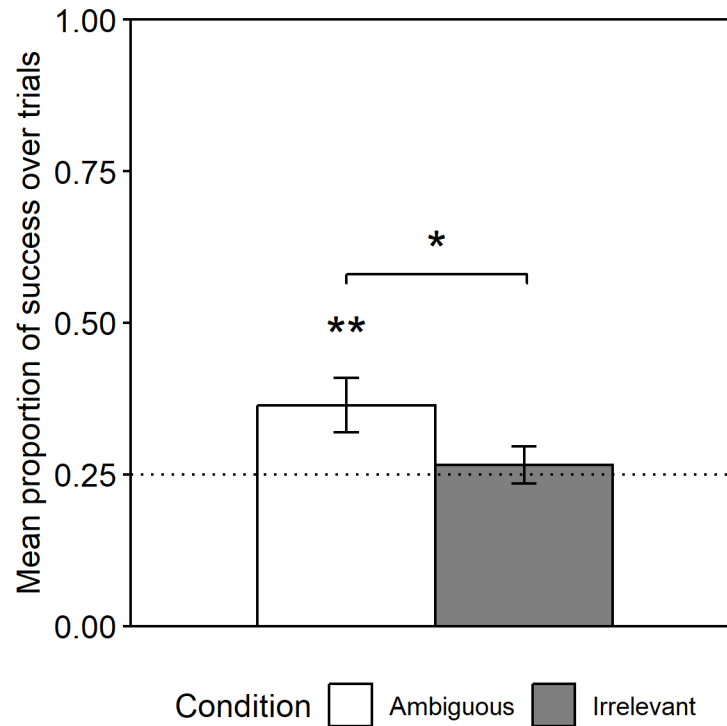
were more likely to select the marked cup than younger ones ($\beta = 0.07$, $SE = 0.03$, $z\text{-value} = 2.00$, $p = .046$).

Next, we performed analyses on the proportion of test trials in which participants selected the box with the unique picture to that predicted by chance — $1/4$, since there were four boxes to choose from. As Figure 6 shows, the average proportion of trials in which the participants first chose the box with the unique picture was significantly higher than that predicted by chance only in the ambiguous cue condition (Ambiguous: $M = 0.36$, $M_{dn} = 0.38$, $SD = 0.19$, $W_+ = 121$, $p = .006$, $r = .78$; Irrelevant: $M = 0.27$, $M_{dn} = 0.25$, $SD = 0.13$, $W_+ = 28$, $p = .55$, $r = .24$, Wilcoxon signed-rank tests).

We also conducted an additional analysis to assess any possible effect of the feature that the cue-card shared with the unique picture in the ambiguous cue condition (i.e., same shape or same pattern). This analysis showed no effect of the feature shared by the cue-card and the unique picture on the participants' performance (for details see additional results and analysis: Study 3 in the Supplementary Materials).

Figure 6.

Performance During the Test Trials of Study 3 Per Condition (Ambiguous and Irrelevant Cue Conditions).



Note: Mean proportion of success in all trials (error bars indicate the SEM). The dotted lines represent the level of performance predicted by chance (1/4). Comparisons of proportions against chance by one-sample Wilcoxon-signed rank tests. Comparisons across conditions by GLMMs (detailed results and model descriptions are reported in the main text).

*: $p < .05$, **: $p < .01$

4.3. Discussion

In Study 3's ambiguous cue condition, toddlers processed the ambiguous novel icon as if it were sufficiently informative to discover the location of the reward. It is likely that the children's choice of the box with the unique picture in the ambiguous cue condition was guided by the cue card. Indeed, in the irrelevant cue condition (which was identical to the informative cue condition, except that the cue cards carried no information about the location of the reward),

the participants did not select the box with the unique picture more often than predicted by chance.

In Study 3, participants managed to find the reward more often than predicted by chance in the ambiguous cue condition. Yet, toddlers' performance remained far from ceiling, perhaps because it required to flexibly alternate between relying on shapes and patterns. Shifting between rules is an executively taxing process that is challenging for young children (Doebel & Zelazo, 2015). Moreover, in Study 3, the participants' performance did not improve significantly across trials. This absence of learning might seem surprising, considering that the participants could have gradually learned that the reward was always located under the box with the unique symbol. This absence of learning might result from the communicative nature of our task. When children are provided with pedagogical information from an adult, they tend to focus on it narrowly, at the expense of other potential sources of information (Bonawitz et al., 2011). Thus, it is possible that in Study 3, participants focused primarily on using the cue-card shown by the experimenter to discover the location of the reward, at the expense at any other strategy that they might have used in a non-communicative context.

5. General Discussion

Our data suggest that toddlers process unfamiliar communicative cues as if they were sufficiently easy to process, informative, and accurate. In all three studies, children had to use a novel cue in order to locate a reward hidden under one of several containers.

Study 1 investigated the effect of processing costs on toddlers' processing of novel communicative cues. In Study 1, the experimenter always placed an unfamiliar communicative cue (a marker) on one of the two opaque cups. In the negative condition, the participants were

explicitly told that the marker was placed where the reward was not. In that case, participants selected the baited cup more often than predicted by chance during their first choice. In short, Study 1's negative condition confirmed that participants could resist reaching for the location where the marker was placed. Conversely, in Study 1's ambiguous condition, the experimenter did not clarify explicitly what the marker conveyed. In that case, the marker could receive at least two possible interpretations which contradicted each other and differed in ease of processing. The marker could be interpreted as directly indicating the location of the cup containing the reward (by being placed on top of it). An alternative and more costly interpretation of the marker was that it indicated where the marker was not, thereby allowing the participants to discover the location of the reward (by excluding the marked cup). Among these two possible interpretations, children were more likely to use the less costly one, despite receiving repeated feedback showing that the marker was always placed on the empty opaque cup. Thus, Study 1 indicate that toddlers' processing of novel cues is guided by an expectation of processing ease.

Studies 2 and 3 provide evidence for the effect of cognitive benefits on toddlers' processing of novel cues. In Study 2, toddlers participated in a hiding game in which a marker was placed on an empty cup. Importantly, assuming that the marker was placed on the empty location was sufficiently informative to discover the location of the reward only in the opaque condition, and not in the transparent condition. Accordingly, participants were more likely to keep searching for the reward under the marked cup in the opaque condition than in the transparent condition. Thus, Study 2's results suggest that toddlers process the novel communicative cue as if it were used in a sufficiently informative manner. In Study 3's ambiguous cue condition, participants had to choose between two interpretations of a picture,

which could be taken to refer either to the pattern or to the shape of the image placed on the container containing the reward. Overall, toddlers were more likely to choose the most informative interpretation of the picture, that is, the interpretation that reduced maximally their uncertainty about the location of the reward. The data of Studies 2-3 dovetail with results showing that preschoolers process familiar communicative cues (labels) by assuming that speakers are informative (Bohn et al., 2022; Frank & Goodman, 2014). Here, we provide evidence for a comparable learning strategy in younger children (toddlers) when processing novel nonverbal cues.

We find that when processing novel communicative cues, toddlers tend to converge towards interpretations that are easy to process, and informative. Importantly, we do not claim that toddlers consider all the possible hypotheses about the interpretation of the novel cues (markers or icons), evaluate their cost-benefit ratios, and subsequently, chose the hypothesis whose cost-benefit ratio is the lowest. Our data can be explained by simpler processes, where toddlers access first the hypothesis that has a better cost-benefit ratio, and never consider alternative hypotheses.

Studies 1-2 also suggest that toddlers expect novel communicative cues to be accurate. Previous studies have shown that, by preschool age, children expect information conveyed by familiar means to be accurate. For instance, young preschoolers find it hard not to trust familiar communicative cues, such as pointing and testimony, even when these familiar cues are misleading (Heyman et al., 2013; Jaswal et al., 2010; Mascaro & Sperber, 2009; Vanderbilt et al., 2011), or when they conflict with other sources of information, such as emotional expressions (Kachel, Hardecker, et al., 2021). Moreover, extensive experience with a communicative cue is not needed to trigger young children's trust. Toddlers rely on unfamiliar

cues (Tomasello et al., 1997; Zlatev et al., 2013), and can trust them very strongly (Mascaro & Kovács, 2022). For instance, in Mascaro and Kovács (2022), an adult used a marker to convey the location of a reward (hidden in one of two cups). First, the adult used the marker reliably, by placing it on top of the reward's location. Next, during the test trials, the participants saw where the reward was hidden, but the adult placed the marker on the empty cup. By 18-months-old onwards, toddlers searched for the reward in the marked location (rather than where they saw it) when the adult was well-informed. These data indicate that very little experience with the reliability of a cue can elicit powerful trust in toddlers.

Our data complement these past studies by showing that toddlers can rely strongly on a novel cue even when they have *never* experienced that this specific cue was reliable in the past. In Mascaro and Kovács (2022), toddlers received some evidence that the novel cue was reliable prior to the test. By contrast, in our Studies 1-2, the novel cue was never placed on the cup containing the reward (neither in the familiarization phase of Study 2, nor in the test phases of Studies 1-2). Yet, toddlers kept searching for the reward under the marked cup in the ambiguous condition of study 1, and in the opaque condition of Study 2. Thus, our data indicate that toddlers can extend their disposition to rely on communicated information to novel cues that they have never experienced to be reliable. Such a capacity is crucial to learn from novel communicative behaviors, such as novel words, or demonstrations.

Note that we do not claim that toddlers' trust in communicated information is unlimited — there is abundant evidence for young children's capacity to select informants, and to assess communicated information from an early age (e.g., Bazhydai et al., 2020; Begus & Southgate, 2012; Castelain et al., 2018; Chow et al., 2008; Dautriche et al., 2021; Kachel, Moore, et al., 2021; Koenig & Echols, 2003; Koenig & Woodward, 2010; Tummeltshammer et al., 2014; for

reviews and meta-analyses, see Clément, 2010; Harris et al., 2018; Poulin-Dubois & Brosseau-Liard, 2016; Sobel & Finiasz, 2020; Tong et al., 2020). Nor do we claim that toddlers process communicated information by assuming that adults are always competent and knowledgeable. Indeed, children modulate their processing of communicated information based on what others have seen from toddlerhood onwards (for a review, see Mazzarella & Pouscoulous, 2021). More simply, our data indicate that when toddlers face a knowledgeable and confident adult communicating using a novel cue, toddlers tend to treat this cue as if it were reliable.

Our findings can be explained by two types of processes (not mutually exclusive). First, they may arise as a consequence of cognitive systems' tendency to be geared toward efficiency. *Ceteris paribus*, minds should be built to prioritize processes that yield important benefits, while reducing the associated costs (Chater & Vitányi, 2003; Sperber & Wilson, 1986). As a result, the cognitive system should focus on simpler hypotheses rather than on more complex ones when processing a novel cue. More generally, a preference for simplicity may characterize cognition at large, such that individuals first generate simpler hypotheses (Chater & Vitányi, 2003). Thus, a preference for simplicity should lead toddlers to consider first, or to give more weight to the hypotheses about the interpretation of cues that are easier to process.

Similarly, cognitive systems might tend to focus on hypotheses whose putative cognitive benefits are high, rather than low, when processing novel cues. We found that toddlers processed novel cues as if they were accurate (Studies 1-2) and informative (Studies 2-3). Some degree of accuracy is a basic property of cues (a completely inaccurate cue is not a cue at all). Likewise, hypothesis generation and hypothesis testing should favor considering and testing first hypotheses that, if proven true, might reduce one's uncertainty. In short, minds should be built to detect cognitively useful regularities before other regularities, and to rely upon those regularities

above others. Cognitively useful regularities yield important benefits while having reduced costs. This general property of minds could explain toddlers' tendency to process novel cues by assuming that their cognitive benefits are high.

Second, toddlers' tendency to process novel communicative cues by assuming that they are cognitively useful may result from communication-specific processes. First, toddlers' positive expectations about novel cues might be triggered by the recognition of an intention to communicate. By offering information, sources of communication imply that the information they provide is cognitively useful (Frank, 2016; Sperber & Wilson, 1986). For this reason, the detection of communicative intentions may trigger specific expectations about the cognitive utility of communicated information. Such expectations should lead learners to process communicative cues as if their cost/benefit ratios were low. Second, the specific source manipulating novel cues might contribute to trigger positive expectations in toddlers. In our studies, cues were handled by seemingly benevolent and competent adults, a priori likely to want to help the child. Thus, toddlers might expect that adults typically provide them with cognitively useful information, with sufficiently high cognitive benefits, and sufficient low processing costs.

Our studies were not designed to adjudicate between the role of general cognitive factors and that of communication-specific factors in shaping toddlers' processing of novel cues. In fact, we find it plausible that both a general tendency to be geared toward cognitive efficiency and communication-specific expectations may jointly contribute to guiding the processing of novel communicative cues. By assuming that communicative cues are easy to process, informative and accurate, learners exert pressure on communicative systems that, in turn, contribute to ensuring that their expectations of cognitive utility are warranted. In the wake of Zipf's (1949) seminal work, many studies suggest that emitters of communicative messages balance the constraint of

minimizing their own efforts with the requirement to fulfill their audience's needs (Caplan et al., 2020; Fedzechkina et al., 2017; Fedzechkina & Jaeger, 2020; Kanwal et al., 2017; Piantadosi et al., 2011; Zipf, 1949). The optimization of communication at the interaction level has cascading consequences on the cultural dynamics at the population level. It is likely to lead to the emergence and stabilization of communicatively efficient systems, which are informative, easy to process, and easy to produce (Fedzechkina et al., 2018; Gibson et al., 2019; Mahowald et al., 2018).

Our data provide direct evidence for the mechanisms that may support the emergence of efficient communicative systems, by demonstrating that learners exert pressure on communicative efficiency. First, we found that toddlers have great difficulties discovering the meaning of cues that are either too costly (because they are difficult to process, in Study 1's ambiguous condition), or not beneficial enough (because they are insufficiently informative, in Study 2's opaque condition). As a result, suboptimal uses of communicative cues are less likely to be understood by toddlers, compared to optimal ones. Second, toddlers readily processed cues as if they were used efficiently. In Study 1's ambiguous condition and Study 2's opaque condition, toddlers processed the novel cue as if it was placed on the cup containing the reward across nine consecutive trials, and despite systematic evidence to the contrary. Thus, toddlers processed novel communicative cues as if they had a low cost/benefit ratio.

Therefore, our data indicate that by toddlerhood, learners shape communicative systems in at least two ways: by discovering the meaning of communicative cues more easily when they are cognitively useful, and by reinterpreting communicative cues whose cognitive utility is suboptimal. Across multiple learners and multiple generations, these two phenomena should give

an advantage to the cultural transmission, diffusion, and stabilization of efficient uses of communicative cues over inefficient ones.

6. References

- Austin, K., Theakston, A., Lieven, E., & Tomasello, M. (2014). Young children's understanding of denial. *Developmental Psychology, 50*, 2061- 2070. <https://doi.org/10.1037/a0037179>
- Bahtiyar, S., & Küntay, A. C. (2009). Integration of communicative partner's visual perspective in patterns of referential requests*. *Journal of Child Language, 36*(3), 529- 555. <https://doi.org/10.1017/S0305000908009094>
- Bannard, C., Rosner, M., & Matthews, D. (2017). What's Worth Talking About? Information Theory Reveals How Children Balance Informativeness and Ease of Production. *Psychological Science, 28*(7), 954- 966. <https://doi.org/10.1177/0956797617699848>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*, 1- 48. <https://doi.org/10.18637/jss.v067.i01>
- Bazhydai, M., Westermann, G., & Parise, E. (2020). "I don't know but I know who to ask": 12-month-olds actively seek information from knowledgeable adults. *Developmental Science, 23*(5), e12938. <https://doi.org/10.1111/desc.12938>
- Begus, K., & Southgate, V. (2012). Infant pointing serves an interrogative function. *Developmental Science, 15*(5), 611- 617. <https://doi.org/10.1111/j.1467-7687.2012.01160.x>

- Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, 8(6), 492- 499. <https://doi.org/10.1111/j.1467-7687.2005.00440.x>
- Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, 109(9), 3253- 3258. <https://doi.org/10.1073/pnas.1113380109>
- Bergen, L., Levy, R., & Goodman, N. (2016). Pragmatic reasoning through semantic inference. *Semantics and Pragmatics*, 9(0), Article 0. <https://doi.org/10.3765/sp.9.20>
- Bohn, M., Tessler, M. H., Merrick, M., & Frank, M. C. (2021). How young children integrate information sources to infer the meaning of words. *Nature Human Behaviour*, 5(8), Article 8. <https://doi.org/10.1038/s41562-021-01145-1>
- Bohn, M., Tessler, M. H., Merrick, M., & Frank, M. C. (2022). Predicting pragmatic cue integration in adults' and children's inferences about novel word meanings. *Journal of Experimental Psychology: General*, 151(11), 2927- 2942. <https://doi.org/10.1037/xge0001216>
- Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, 120(3), 322- 330. <https://doi.org/10.1016/j.cognition.2010.10.001>
- Burks, A. W. (1949). Icon, Index, and Symbol. *Philosophy and Phenomenological Research*, 9(4), 673- 689. <https://doi.org/10.2307/2103298>
- Caplan, S., Kodner, J., & Yang, C. (2020). Miller's monkey updated : Communicative efficiency and the statistics of words in natural language. *Cognition*, 205, 104466. <https://doi.org/10.1016/j.cognition.2020.104466>

- Castelain, T., Bernard, S., & Mercier, H. (2018). Evidence that Two-Year-Old Children are Sensitive to Information Presented in Arguments. *Infancy*, 23(1), 124- 135. <https://doi.org/10.1111/infa.12202>
- Chater, N., & Vitányi, P. (2003). Simplicity : A unifying principle in cognitive science? *Trends in Cognitive Sciences*, 7(1), 19- 22. [https://doi.org/10.1016/S1364-6613\(02\)00005-0](https://doi.org/10.1016/S1364-6613(02)00005-0)
- Chow, V., Poulin-Dubois, D., & Lewis, J. (2008). To see or not to see : Infants prefer to follow the gaze of a reliable looker. *Developmental Science*, 11(5), 761- 770. <https://doi.org/10.1111/j.1467-7687.2008.00726.x>
- Clark, H. H. (1996). *Using language*. Cambridge University Press.
- Clément, F. (2010). To Trust or not to Trust? Children's Social Epistemology. *Review of Philosophy and Psychology*, 1(4), 531- 549. <https://doi.org/10.1007/s13164-010-0022-3>
- Couillard, N. L., & Woodward, A. L. (1999). Children's comprehension of deceptive points. *British Journal of Developmental Psychology*, 17(4), 515- 521. <https://doi.org/10.1348/026151099165447>
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13(4), 148- 153. <https://doi.org/10.1016/j.tics.2009.01.005>
- Dautriche, I., Goupil, L., Smith, K., & Rabagliati, H. (2021). Knowing How You Know : Toddlers Reevaluate Words Learned From an Unreliable Speaker. *Open Mind*, 5, 1- 19. https://doi.org/10.1162/opmi_a_00038
- Davies, C., & Katsos, N. (2010). Over-informative children : Production/comprehension asymmetry or tolerance to pragmatic violations? *Lingua*, 120(8), 1956- 1972. <https://doi.org/10.1016/j.lingua.2010.02.005>

- Degen, J., Franke, M., & Jager, G. (2013). Cost-Based Pragmatic Inference about Referential Expressions. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 35(35), 6.
- Degen, J., & Tanenhaus, M. K. (2019). Constraint-Based Pragmatic Processing. In C. Cummins & N. Katsos (Éds.), *The Oxford Handbook of Experimental Semantics and Pragmatics* (p. 20- 38). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198791768.013.8>
- Diesendruck, G., & Bloom, P. (2003). How Specific is the Shape Bias? *Child Development*, 74(1), 168- 178. <https://doi.org/10.1111/1467-8624.00528>
- Doebel, S., & Zelazo, P. D. (2015). A meta-analysis of the Dimensional Change Card Sort : Implications for developmental theories and the measurement of executive function in children. *Developmental Review*, 38, 241- 268. <https://doi.org/10.1016/j.dr.2015.09.001>
- Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., Marjanovič-Umek, L., Gayraud, F., Kovacevic, M., & Gallego, C. (2012). Differences between girls and boys in emerging language skills: Evidence from 10 language communities. *British Journal of Developmental Psychology*, 30(2), 326- 343. <https://doi.org/10.1111/j.2044-835X.2011.02042.x>
- Eskritt, M., Whalen, J., & Lee, K. (2008). Preschoolers can recognize violations of the Gricean maxims. *British Journal of Developmental Psychology*, 26(3), 435- 443. <https://doi.org/10.1348/026151007X253260>
- Falkum, I. L., Recasens, M., & Clark, E. V. (2017). “The moustache sits down first” : On the acquisition of metonymy*. *Journal of Child Language*, 44(1), 87- 119. <https://doi.org/10.1017/S0305000915000720>

-
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3 : A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175- 191. <https://doi.org/10.3758/BF03193146>
- Fedzechkina, M., Chu, B., & Florian Jaeger, T. (2018). Human Information Processing Shapes Language Change. *Psychological Science*, *29*(1), 72- 82. <https://doi.org/10.1177/0956797617728726>
- Fedzechkina, M., & Jaeger, T. F. (2020). Production efficiency can cause grammatical change : Learners deviate from the input to better balance efficiency against robust message transmission. *Cognition*, *196*, 104115. <https://doi.org/10.1016/j.cognition.2019.104115>
- Fedzechkina, M., Newport, E. L., & Jaeger, T. F. (2017). Balancing Effort and Information Transmission During Language Acquisition : Evidence From Word Order and Case Marking. *Cognitive Science*, *41*(2), 416- 446. <https://doi.org/10.1111/cogs.12346>
- Feiman, R., Mody, S., Sanborn, S., & Carey, S. (2017). What Do You Mean, No? Toddlers' Comprehension of Logical “No” and “Not”. *Language Learning and Development*, *13*(4), 430- 450. <https://doi.org/10.1080/15475441.2017.1317253>
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development*, *59*(5), i- 185. <https://doi.org/10.2307/1166093>
- Frank, M. C. (2016). *Rational speech act models of pragmatic reasoning in reference games*. PsyArXiv. <https://doi.org/10.31234/osf.io/f9y6b>
- Frank, M. C., & Goodman, N. D. (2012). Predicting Pragmatic Reasoning in Language Games. *Science*, *336*(6084), 998- 998. <https://doi.org/10.1126/science.1218633>

-
- Frank, M. C., & Goodman, N. D. (2014). Inferring word meanings by assuming that speakers are informative. *Cognitive Psychology*, 75, 80- 96.
<https://doi.org/10.1016/j.cogpsych.2014.08.002>
- Franke, M. (2011). Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantics and Pragmatics*, 4(0), Article 0. <https://doi.org/10.3765/sp.4.1>
- Gautam, S., Suddendorf, T., & Redshaw, J. (2021). When can young children reason about an exclusive disjunction? A follow up to Mody and Carey (2016). *Cognition*, 207, 104507.
<https://doi.org/10.1016/j.cognition.2020.104507>
- Gibbs, R. W., & Bryant, G. A. (2008). Striving for optimal relevance when answering questions. *Cognition*, 106(1), 345- 369. <https://doi.org/10.1016/j.cognition.2007.02.008>
- Gibson, E., Futrell, R., Piantadosi, S. P., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How Efficiency Shapes Human Language. *Trends in Cognitive Sciences*, 23(5), 389- 407. <https://doi.org/10.1016/j.tics.2019.02.003>
- Grice, H. P. (1975). Logic and Conversation. In P. Cole & J. L. Morgan (Éds.), *Speech Acts* (p. 41- 58). BRILL. https://doi.org/10.1163/9789004368811_003
- Grosse, G., Moll, H., & Tomasello, M. (2010). 21-Month-olds understand the cooperative logic of requests. *Journal of Pragmatics*, 42(12), 3377- 3383.
<https://doi.org/10.1016/j.pragma.2010.05.005>
- Grosse, G., & Tomasello, M. (2012). Two-year-old children differentiate test questions from genuine questions*. *Journal of Child Language*, 39(1), 192- 204.
<https://doi.org/10.1017/S0305000910000760>

- Gweon, H., & Asaba, M. (2018). Order Matters : Children's Evaluation of Underinformative Teachers Depends on Context. *Child Development*, 89(3), e278- e292. <https://doi.org/10.1111/cdev.12825>
- Gweon, H., Shafto, P., & Schulz, L. (2018). Development of children's sensitivity to overinformativeness in learning and teaching. *Developmental Psychology*, 54(11), 2113- 2125. <https://doi.org/10.1037/dev0000580>
- Harris, P. L. (2012). Trusting What You're Told : How Children Learn from Others. In *Trusting What You're Told*. Harvard University Press. <https://doi.org/10.4159/harvard.9780674065192>
- Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive Foundations of Learning from Testimony. *Annual Review of Psychology*, 69(1), 251- 273. <https://doi.org/10.1146/annurev-psych-122216-011710>
- Heyman, G. D., Sritanyaratana, L., & Vanderbilt, K. E. (2013). Young Children's Trust in Overtly Misleading Advice. *Cognitive Science*, 37(4), 646- 667. <https://doi.org/10.1111/cogs.12020>
- Hill, A., Collier-Baker, E., & Suddendorf, T. (2012). Inferential reasoning by exclusion in children (Homo sapiens). *Journal of Comparative Psychology*, 126(3), 243- 254. <https://doi.org/10.1037/a0024449>
- Jaswal, V. K., Croft, A. C., Setia, A. R., & Cole, C. A. (2010). Young Children Have a Specific, Highly Robust Bias to Trust Testimony. *Psychological Science*, 21(10), 1541- 1547. <https://doi.org/10.1177/0956797610383438>

- Kachel, G., Hardecker, D. J. K., & Bohn, M. (2021). Young children's developing ability to integrate gestural and emotional cues. *Journal of Experimental Child Psychology*, *201*, 104984. <https://doi.org/10.1016/j.jecp.2020.104984>
- Kachel, G., Moore, R., Hepach, R., & Tomasello, M. (2021). Toddlers Prefer Adults as Informants: 2- and 3-Year-Olds' Use of and Attention to Pointing Gestures From Peer and Adult Partners. *Child Development*, *92*(4), e635- e652. <https://doi.org/10.1111/cdev.13544>
- Kanwal, J., Smith, K., Culbertson, J., & Kirby, S. (2017). Zipf's Law of Abbreviation and the Principle of Least Effort: Language users optimise a miniature lexicon for efficient communication. *Cognition*, *165*, 45- 52. <https://doi.org/10.1016/j.cognition.2017.05.001>
- Köder, F., & Falkum, I. L. (2020). Children's metonymy comprehension: Evidence from eye-tracking and picture selection. *Journal of Pragmatics*, *156*, 191- 205. <https://doi.org/10.1016/j.pragma.2019.07.007>
- Koenig, M. A., & Echols, C. H. (2003). Infants' understanding of false labeling events: The referential roles of words and the speakers who use them. *Cognition*, *87*(3), 179- 208. [https://doi.org/10.1016/S0010-0277\(03\)00002-7](https://doi.org/10.1016/S0010-0277(03)00002-7)
- Koenig, M. A., & Woodward, A. L. (2010). Sensitivity of 24-month-olds to the prior inaccuracy of the source: Possible mechanisms. *Developmental Psychology*, *46*(4), 815- 826. <https://doi.org/10.1037/a0019664>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1- 26. <https://doi.org/10.18637/jss.v082.i13>

- Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, 3(3), 299- 321. [https://doi.org/10.1016/0885-2014\(88\)90014-7](https://doi.org/10.1016/0885-2014(88)90014-7)
- Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition*, 108(3), 732- 739. <https://doi.org/10.1016/j.cognition.2008.06.013>
- Mahowald, K., Dautriche, I., Gibson, E., & Piantadosi, S. T. (2018). Word Forms Are Structured for Efficient Use. *Cognitive Science*, 42(8), 3116- 3134. <https://doi.org/10.1111/cogs.12689>
- Mangiafico, S. (2022). *rcompanion: Functions to Support Extension Education Program Evaluation* (2.4.15) [Logiciel]. <https://CRAN.R-project.org/package=rcompanion>
- Mascaro, O., & Kovács, Á. M. (2022). The origins of trust : Humans' reliance on communicative cues supersedes firsthand experience during the second year of life. *Developmental Science*, 00(e13223). <https://doi.org/10.1111/desc.13223>
- Mascaro, O., & Morin, O. (2015). Epistemology for Beginners : Two- to Five-Year-Old Children's Representation of Falsity. *PLOS ONE*, 10(10), e0140658. <https://doi.org/10.1371/journal.pone.0140658>
- Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition*, 112(3), 367- 380. <https://doi.org/10.1016/j.cognition.2009.05.012>
- Mascaro, O., & Sperber, D. (2019). The pragmatic role of trust in young children's interpretation of unfamiliar signals. *PLOS ONE*, 14(10), e0224648. <https://doi.org/10.1371/journal.pone.0224648>

- Matthews, D., Butcher, J., Lieven, E., & Tomasello, M. (2012). Two- and Four-Year-Olds Learn to Adapt Referring Expressions to Context: Effects of Distracters and Feedback on Referential Communication. *Topics in Cognitive Science*, 4(2), 184- 210. <https://doi.org/10.1111/j.1756-8765.2012.01181.x>
- Matthews, D., Lieven, E., & Tomasello, M. (2007). How Toddlers and Preschoolers Learn to Uniquely Identify Referents for Others : A Training Study. *Child Development*, 78(6), 1744- 1759. <https://doi.org/10.1111/j.1467-8624.2007.01098.x>
- Mazzarella, D., & Pouscoulous, N. (2021). Pragmatics and epistemic vigilance : A developmental perspective. *Mind & Language*, 36(3), 355- 376. <https://doi.org/10.1111/mila.12287>
- Meng, X., & Hashiya, K. (2014). Pointing Behavior in Infants Reflects the Communication Partner's Attentional and Knowledge States : A Possible Case of Spontaneous Informing. *PLOS ONE*, 9(9), e107579. <https://doi.org/10.1371/journal.pone.0107579>
- Mody, S., & Carey, S. (2016). The emergence of reasoning by the disjunctive syllogism in early childhood. *Cognition*, 154, 40- 48. <https://doi.org/10.1016/j.cognition.2016.05.012>
- Morisseau, T., Davies, C., & Matthews, D. (2013). How do 3- and 5-year-olds respond to under- and over-informative utterances? *Journal of Pragmatics*, 59, 26- 39. <https://doi.org/10.1016/j.pragma.2013.03.007>
- Nilsen, E. S., & Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology*, 58(2), 220- 249. <https://doi.org/10.1016/j.cogpsych.2008.07.002>
- Nordmeyer, A. E., & Frank, M. C. (2018). Early Understanding of Pragmatic Principles in Children's Judgments of Negative Sentences. *Language Learning and Development*, 14(4), 262- 278. <https://doi.org/10.1080/15475441.2018.1463850>

- O'Neill, D. K. (1996). Two-Year-Old Children's Sensitivity to a Parent's Knowledge State When Making Requests. *Child Development*, 67(2), 659- 677. <https://doi.org/10.2307/1131839>
- O'Neill, D. K., & Topolovec, J. C. (2001). Two-year-old children's sensitivity to the referential (in)efficacy of their own pointing gestures. *Journal of Child Language*, 28(1), 1- 28. <https://doi.org/10.1017/S0305000900004566>
- Palmquist, C. M., Kondrad, R. L., & Norris, M. N. (2018). Follow my point? Preschoolers' expectations about veridicality disrupt their understanding of deceptive points. *Cognitive Development*, 48, 190- 202. <https://doi.org/10.1016/j.cogdev.2018.08.009>
- Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, 108(9), 3526- 3529. <https://doi.org/10.1073/pnas.1012551108>
- Posit Team. (2023). *Integrated Development Environment for R* (2023.3.0.386) [Logiciel]. Posit Software, PBC. <http://www.posit.co/>
- Poulin-Dubois, D., & Brosseau-Liard, P. (2016). The Developmental Origins of Selective Social Learning. *Current Directions in Psychological Science*, 25(1), 60- 64. <https://doi.org/10.1177/0963721415613962>
- Pouscoulous, N., & Tomasello, M. (2020). Early birds : Metaphor understanding in 3-year-olds. *Journal of Pragmatics*, 156, 160- 167. <https://doi.org/10.1016/j.pragma.2019.05.021>
- Quine, W. V. O. (1960). *Word and Object*. MIT Press.
- R Core Team. (2022). *R: A Language and Environment for Statistical Computing* [Logiciel]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rohde, H., Seyfarth, S., Clark, B., Jaeger, G., & Kaufmann, S. (2012). Communicating with Cost-based Implicature : A Game-Theoretic Approach to Ambiguity. *Proceedings of*

-
- SemDial 2012 (SeineDial): The 16th Workshop on Semantics and Pragmatics of Dialogue*, 107- 116. <https://www.research.ed.ac.uk/en/publications/communicating-with-cost-based-implicature-a-game-theoretic-approa>
- Schulze, C., Grassmann, S., & Tomasello, M. (2013). 3-Year-Old Children Make Relevance Inferences in Indirect Verbal Communication. *Child Development*, 84(6), 2079- 2093. <https://doi.org/10.1111/cdev.12093>
- Schulze, C., & Tomasello, M. (2015). 18-month-olds comprehend indirect communicative acts. *Cognition*, 136, 91- 98. <https://doi.org/10.1016/j.cognition.2014.11.036>
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379- 423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Sobel, D. M., & Finiasz, Z. (2020). How Children Learn From Others : An Analysis of Selective Word Learning. *Child Development*, 91(6), e1134- e1161. <https://doi.org/10.1111/cdev.13415>
- Sperber, D., & Wilson, D. (1986). *Relevance : Communication and cognition*. Harvard University Press.
- Sperber, D., & Wilson, D. (2002). Pragmatics, Modularity and Mind-reading. *Mind & Language*, 17(1- 2), 3- 23. <https://doi.org/10.1111/1468-0017.00186>
- Stephens, G., & Matthews, D. (2014). The communicative infant from 0-18 months : The social-cognitive foundations of pragmatic development. In *Pragmatic development in first language acquisition* (p. 13- 35). John Benjamins Publishing Company. <https://doi.org/10.1075/tilar.10.02ste>

- Stiller, A. J., Goodman, N. D., & Frank, M. C. (2015). Ad-hoc Implicature in Preschool Children. *Language Learning and Development, 11*(2), 176- 190. <https://doi.org/10.1080/15475441.2014.927328>
- Szabó, E., & Kovács, Á.-M. (2022). *Infants' early understanding of different forms of negation*. PsyArXiv. <https://doi.org/10.31234/osf.io/nfv4j>
- Tomasello, M. (2010). *Origins of human communication* (p. xiii, 393). MIT Press.
- Tomasello, M., Call, J., & Gluckman, A. (1997). Comprehension of Novel Communicative Signs by Apes and Human Children. *Child Development, 68*(6), 1067- 1080. <https://doi.org/10.2307/1132292>
- Tong, Y., Wang, F., & Danovitch, J. (2020). The role of epistemic and social characteristics in children's selective trust: Three meta-analyses. *Developmental Science, 23*(2), e12895. <https://doi.org/10.1111/desc.12895>
- Tummeltshammer, K. S., Wu, R., Sobel, D. M., & Kirkham, N. Z. (2014). Infants Track the Reliability of Potential Informants. *Psychological Science, 25*(9), 1730- 1738. <https://doi.org/10.1177/0956797614540178>
- Van Der Henst, J., Carles, L., & Sperber, D. (2002). Truthfulness and Relevance in Telling The Time. *Mind & Language, 17*(5), 457- 466. <https://doi.org/10.1111/1468-0017.00207>
- Vanderbilt, K. E., Liu, D., & Heyman, G. D. (2011). The Development of Distrust. *Child Development, 82*(5), 1372- 1380. <https://doi.org/10.1111/j.1467-8624.2011.01629.x>
- Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A Grammar of Data Manipulation* (v. 1.1.0) [Logiciel]. <https://CRAN.R-project.org/package=dplyr>
- Wilson, D., & Sperber, D. (2006). Relevance Theory. In *The Handbook of Pragmatics* (p. 606- 632). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470756959.ch27>

Zipf, G. K. (1949). *Human behavior and the principle of least effort* (p. xi, 573). Addison-Wesley Press.

Zlatev, J., Madsen, E. A., Lenninger, S., Persson, T., Sayehli, S., Sonesson, G., & van de Weijer, J. (2013). Understanding communicative intentions and semiotic vehicles by children and chimpanzees. *Cognitive Development*, 28(3), 312- 329.
<https://doi.org/10.1016/j.cogdev.2013.05.001>

7. Supplementary Materials

7.1. Demographic Information About the Recruitment Pool

Our participants were recruited from large French urban areas (> 1000000 inhabitants). Below, we report official data about the living conditions in our recruitment area (extracted in February 2023 from the repository of the National Institute of Statistics and Economics Studies, <https://www.insee.fr/fr/statistiques/5039899?sommaire=5040030>). To make the characterization of our recruitment area more informative, we also report information about French national averages. The median annual income per household is slightly above the national average (Lyon: 22510€; Paris: 23210€; French average: 21110€). In our recruitment area, more than 50% of the out-of-school population has a college degree or equivalent (Lyon: 55%; Paris: 63%; French average: 31%). The poverty and unemployment rates of our recruitment area are respectively 15% and 12% (both in Lyon and Paris); they are comparable to the national averages (French average poverty rate: 15%; French average unemployment rate: 13%). In our recruitment area, more than 2/3rd of children lives in a so-called "traditional" family, where all the children living in the home are those of the couple, whereas less than 20% of children live in a single-parent family (French average number of single parent families: 23%). In short, in comparison to the national average, the population in our recruitment pool is more educated, and slightly more affluent.

The daycare centers that participated in the study were private. Their cost is state-regulated and the same as the cost of public daycare centers (approximately from 0.15€/h to 4€/h, depending on families' income level and number of children — families with less resources pay less).

7.2. Additional Results and Analyses: Study 1

7.2.1. Study 1's Warm-up Phase

During Study 1's warm-up phase, participants invariably selected the baited cup. Thus, the proportion of warm-up trials in which participants selected the baited cup was higher than predicted by chance (1/3), in both conditions (Ambiguous: $M = 1$, $M_{dn} = 1$, $SD = 0$, $W_+ = 171$, $p < .001$, $r = 1$; Negative: $M = 1$, $M_{dn} = 1$, $SD = 0$, $W_+ = 171$, $p < .001$, $r = 1$; Wilcoxon signed-rank tests). Therefore, during warm-up trials participants selected one the two empty cups less often than predicted by chance, in both conditions ($M = 0$, $M_{dn} = 0$, $SD = 0$, $W_+ = 0$, $p < .001$, $r = -1$ for the empty opaque and transparent cups in both conditions, Wilcoxon signed-rank tests). In short, participants were at ceiling in selecting the baited cup during Study 1's warm-up trials — thus establishing that they were motivated to find the reward.

7.2.2. Study 1's Test Phase: Second Choice of Cup

When analyzing the participants' second choice of cup, we did not run GLMMs because the participants sometimes found the pompom on their first attempt, thus resulting in missing data for their second choice in multiple trials.

Baited Cup. Participants selected the baited cup as their second choice significantly more often than predicted by chance in both condition (Ambiguous: $M = 0.85$, $M_{dn} = 0.86$, $SD = 0.14$, $W_+ = 91$, $p < .001$, $r = 1$; Negative: $M = 0.95$, $M_{dn} = 1$, $SD = 0.16$, $W_+ = 171$, $p < .001$, $r = 1$; Wilcoxon signed-rank tests).

Marked Cup. The mean proportion of trials in which the participants selected the marked cup on their second attempt was significantly lower than that predicted by chance in both conditions (Ambiguous: $M = 0.11$, $M_{dn} = 0.06$, $SD = 0.12$, $W_+ = 0$, $p < .001$, $r = -1$; Negative: M

= 0, $M_{dn} = 0$, $SD = 0$, $W_+ = 0$, $p < .001$, $r = -1$ Wilcoxon signed-rank tests). Thus, Study 1's participants had no difficulty excluding the marked cup once they had discovered that it was empty.

Transparent Cup. During their second choice the participants selected the transparent cup less often than predicted by chance in both conditions (Ambiguous: $M = 0.05$, $M_{dn} = 0$, $SD = 0.09$, $W_+ = 0$, $p < .001$, $r = -1$; Negative: $M = 0.05$, $M_{dn} = 0$, $SD = 0.16$, $W_+ = 2$, $p < .001$, $r = -.96$, Wilcoxon signed-rank tests).

7.3. Additional Results and Analyses: Study 2

7.3.1. Study 2's Warm-up Phase

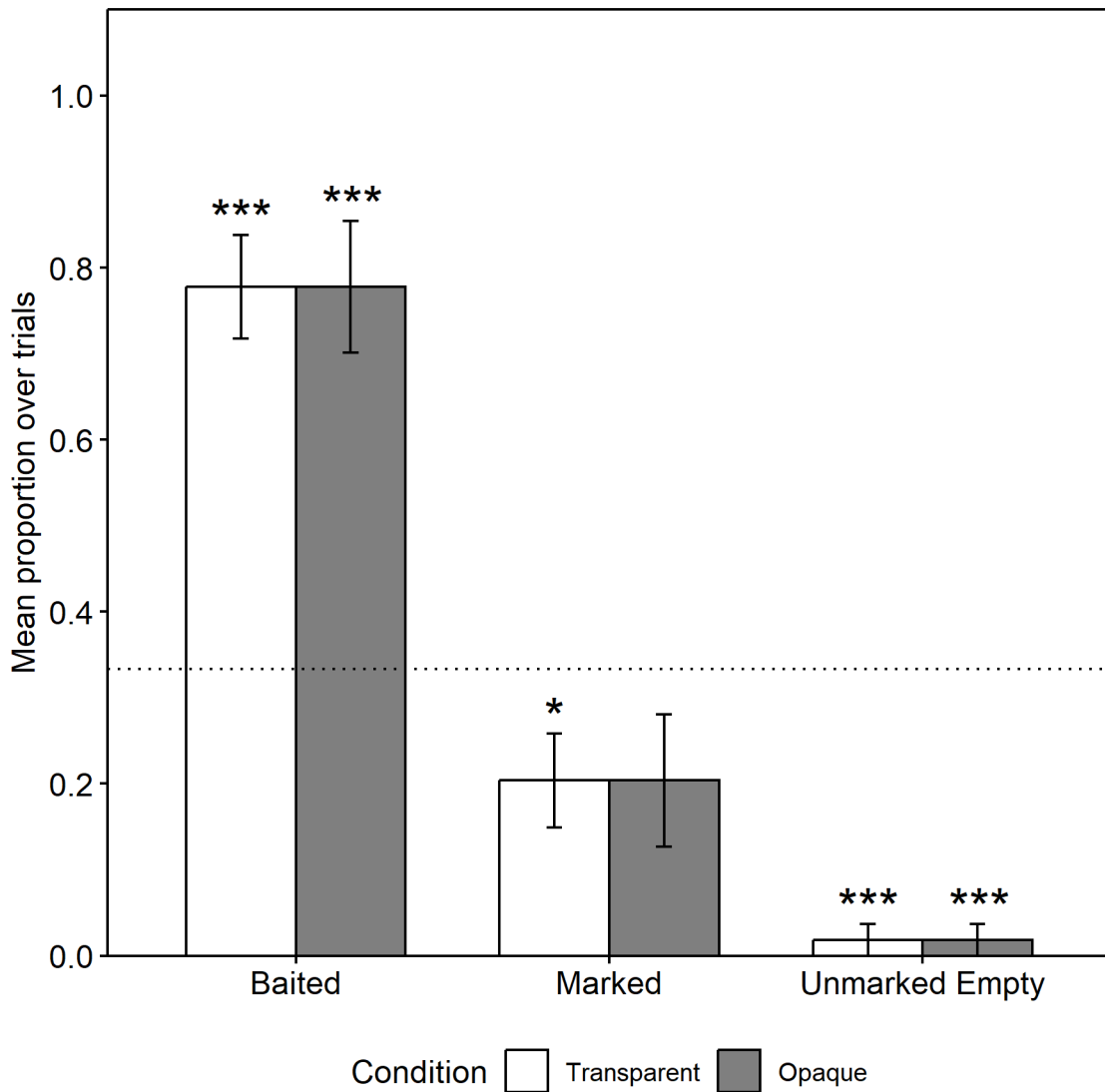
During Study 2's warm-up phase, participants almost always selected the baited cup. The proportion of warm-up trials in which participants selected the baited cup was higher than predicted by chance (1/3), in both conditions (Transparent: $M = 0.94$, $M_{dn} = 1$, $SD = 0.13$, $W_+ = 171$, $p < .001$, $r = 1$; Opaque: $M = 1$, $M_{dn} = 1$, $SD = 0$, $W_+ = 171$, $p < .001$, $r = 1$, Wilcoxon signed-rank tests). For Study 2's warm-up trials, the proportion of trials in which participants selected one the two empty cups was lower than predicted by chance, in both conditions (Transparent: empty opaque cup: $M = 0.06$, $M_{dn} = 0$, $SD = 0.13$, $W_+ = 0$, $p < .001$, $r = -1$; empty transparent cup: $M = 0$, $M_{dn} = 0$, $SD = 0$, $W_+ = 0$, $p < .001$, $r = -1$; Opaque: first empty opaque cup: $M = 0$, $M_{dn} = 0$, $SD = 0$, $W_+ = 0$, $p < .001$, $r = -1$; second empty opaque cup: $M = 0$, $M_{dn} = 0$, $SD = 0$, $W_+ = 0$, $p < .001$, $r = -1$; Wilcoxon signed-rank tests). In short, participants systematically selected the baited cup during Study 2's warm-up trials.

7.3.2. Study 2's Familiarization Phase

As Figure S1 shows, the average proportion of familiarization trials in which participants selected the baited cup was higher than predicted by chance in both conditions (Figure S1; Transparent: $M = 0.78$, $M_{dn} = 0.83$, $SD = 0.26$, $W_+ = 165$, $p < .001$, $r = .93$; Opaque: $M = 0.78$, $M_{dn} = 1$, $SD = 0.32$, $W_+ = 161$, $p < .001$, $r = .88$, Wilcoxon signed-rank tests). The average proportion of familiarization trials in which participants selected the marked cup was lower than predicted by chance in the Transparent condition (Transparent: $M = 0.20$, $M_{dn} = 0.17$, $SD = 0.23$, $W_+ = 35$, $p = .026$, $r = -.59$) and it did not differ from chance in the Opaque condition ($M = 0.20$, $M_{dn} = 0$, $SD = 0.33$, $W_+ = 66$, $p = .39$, $r = -0.23$, Wilcoxon signed-rank tests). The average proportion of familiarization trials in which participants selected the unmarked empty cup was lower than predicted by chance, in both conditions (Figure S1; Transparent: $M = 0.02$, $M_{dn} = 0$, $SD = 0.08$, $W_+ = 0$, $p < .001$, $r = -1$; Opaque: $M = 0.02$, $M_{dn} = 0$, $SD = 0.08$, $W_+ = 0$, $p < .001$, $r = -1$; Wilcoxon signed-rank tests). In short, participants succeeded in selecting the baited cup more often than predicted by chance during Study 2's familiarization trials.

Figure S1

Proportion Selection for Each Cup in Study 2's Familiarization Phase.



Note: Average proportion of trials in which participants selected each cup (baited, empty, marked) per condition (ambiguous vs. negative; transparent vs. opaque) during Study 2's familiarization trials. Error bars represent standard errors. The dotted lines represent the level of performance predicted by chance (i.e., 1/3). Comparisons against chance by Wilcoxon signed-rank tests.

*: $p < .05$, ***: $p < .001$.

7.3.3. Study 2's test phase: Second choice of cup.

As in Study 1, we did not run GLMMs on the participants' second choice of cup because of missing data for the trials whose participant found the reward during their first choice.

Baited Cup. In Study 2, the proportion of trials in which participants selected the baited cup as their second choice was significantly higher than predicted by chance in both conditions (Transparent: $M = 0.91$, $M_{dn} = 1$, $SD = 0.15$, $W_+ = 91$, $p = .001$, $r = 1$; Opaque: $M = 0.49$, $M_{dn} = 0.5$, $SD = 0.23$, $W_+ = 145$, $p = .010$, $r = .70$; Wilcoxon signed-rank tests).

Marked Cup. For the second choice, the participants selected the marked cup significantly less often than predicted by chance in both conditions (Transparent: $M = 0.01$, $M_{dn} = 0$, $SD = 0.03$, $W_+ = 0$, $p < .001$, $r = -1$; Opaque: $M = 0.09$, $M_{dn} = 0.11$, $SD = 0.10$, $W_+ = 0$, $p < .001$, $r = -.1$; Wilcoxon signed-rank tests), thus confirming that participants could exclude the marked cup once they had discovered that it was empty.

Unmarked Empty Cup. For their second choice, toddlers selected the unmarked empty cup significantly less often than predicted by chance in the transparent but not in the opaque condition (Transparent: $M = 0.08$, $M_{dn} = 0$, $SD = 0.15$, $W_+ = 3$, $p = .002$, $r = -.93$; Opaque: $M = 0.42$, $M_{dn} = 0.40$, $SD = 0.23$, $W_+ = 118$, $p = .16$, $r = .38$; Wilcoxon signed-rank tests).

7.3.4. Comparison of Performance: Study 1 vs. 2

Study 2's transparent condition was modeled after Study 1's ambiguous condition, with the addition of a familiarization phase designed to increase participants' performance. In an exploratory analysis, we checked whether adding a familiarization phase to Study 1's ambiguous condition increased participants' tendency to select the baited cup, and reduced their tendency to

select the marked cup. We assessed the effect of adding a familiarization phase by running GLMMs on the participants' first choice of a given cup as a binary repeated dependent measure. We fitted (1) null models, which included subject identity as a random intercept, and the fixed effect of Age (in months), Trial number (1–9) and Sex (female vs. male) and (2) full models, which added the fixed effect of condition (Study 1's ambiguous condition vs. Study 2's transparent condition) to the null models.

Test Phase, First Choice of Baited Cup. Adding a familiarization phase to Study 1's ambiguous condition increased participants' tendency to select first the baited cup during the test phase. The full model indicated that participants' proportion of first choices of the baited cup during test trials was significantly higher in Study 2's transparent condition than in Study 1's ambiguous condition ($\beta = 2.53$, $SE = 0.76$, $z\text{-value} = 3.31$, $p < .001$, LRT comparing null vs. full models: $\chi^2(1) = 10.99$, $p < .001$). The full model revealed no other significant effect.

Test Phase, First Choice of Marked Cup. Participants' tendency to select first the marked cup was significantly lower in Study 2's transparent condition than in Study 1's ambiguous condition ($\beta = -2.30$, $SE = 0.75$, $z\text{-value} = -3.06$, $p = .002$; LRT comparing null vs. full models: $\chi^2(1) = 9.29$, $p = .002$). The full model also showed a significant main effect of Trial number, indicating that the participants' tendency to select first the marked cup decreased across trials ($\beta = -0.14$, $SE = 0.06$, $z\text{-value} = -2.34$, $p = .019$). The full model revealed no other significant effect.

Test Phase, First Choice of Transparent Cup. Participants' tendency to select the transparent cup did not differ significantly across conditions (Study 1's ambiguous condition vs. Study 2's transparent condition, LRT comparing null vs. full models: $\chi^2(1) = .36$, $p = .55$). The full model ran on participants' first choices of the transparent cup revealed no significant effect.

In sum, the results confirmed that, during the test phase, participants were less likely to select first the baited cup and more likely to select first the marked cup in Study 1's ambiguous condition than in Study 2's transparent condition. Thus, as expected, adding a familiarization phase to Study 1's ambiguous condition improved participants' performance during the test phase.

7.3.5. Complementary Analysis: Reanalysis of Study 2's Data Including the Supernumerary Participants

Due to a recruitment error in the opaque condition of Study 2, four additional participants were tested beyond the 18 planned. These participants were excluded from the analysis reported in the main text. Yet, we also checked whether our results would remain the same after including these four supernumerary participants. To this end, we reanalyzed data from the test phase (first and second choice), while including all the participants tested (Transparent: $n = 18$, $M = 29.2$ months, $range = 23\text{--}35$ months, $SD = 2.92$; Opaque: $n = 22$, $M = 29.1$ months, $range = 23\text{--}37$ months, $SD = 3.73$). This complementary analysis confirmed that the pattern of statistically significant results remained the same irrespective of whether the super-numeracy participants were included in the analyses.

7.3.5.1. Test phase: First choice

Baited Cup. The full model ran on data including the supernumerary participants confirmed that toddlers selected the baited cup significantly more often in the transparent condition than in the opaque condition ($\beta = 2.92$, $SE = 0.59$, $z\text{-value} = 4.96$, $p < .001$, LRT comparing null vs. full models: $\chi^2(1) = 22.16$, $p < .001$). In addition, the full model revealed a main effect of Age indicating older participants were more likely to select the baited cup than younger ones during their first choices ($\beta = 0.21$, $SE = 0.08$, $z\text{-value} = 2.51$, $p = .012$). In the

dataset including the supernumerary participants, toddlers selected the baited cup as their first choice less often than predicted by chance during the test trials of the opaque condition ($M = 0.15$, $M_{dn} = 0.12$, $SD = 0.15$, $W_+ = 17$, $p < .001$, $r = -.87$, Wilcoxon signed-rank test).

Marked Cup. The full model ran on data including the supernumerary participants confirmed that participants were significantly less likely to select the marked cup during their first choice in the opaque condition than in the transparent condition ($\beta = -2.22$, $SE = 0.76$, $z\text{-value} = -2.91$, $p = .004$; LRT comparing null vs. full models: $\chi^2(1) = 8.54$, $p = .003$). In addition, the full model revealed a main effect of Age indicating older participants were less likely to select the marked cup than younger ones ($\beta = -0.30$, $SE = 0.11$, $z\text{-value} = -2.67$, $p = .008$).

In the dataset including the supernumerary participants, toddlers selected the marked cup during their first choice, more often than predicted by chance in the opaque condition ($M = 0.66$, $M_{dn} = 0.76$, $SD = 0.33$, $W_+ = 228$, $p = .001$, $r = .80$, Wilcoxon signed-rank test).

Unmarked Empty Cup. During the test phase of Study 2's opaque condition the participants also selected the unmarked empty cup less often than predicted by chance ($M = 0.19$, $M_{dn} = 0.12$, $SD = 0.20$, $W_+ = 46$, $p = .009$, $r = -.64$, Wilcoxon signed-rank test).

7.3.5.2. Test Phase: Second Choice of Cup

Baited Cup. In the dataset including the supernumerary participants, the proportion of trials whose participant selected the baited cup as their second choice was significantly higher than predicted by chance in the opaque condition ($M = 0.49$, $M_{dn} = 0.47$, $SD = 0.25$, $W_+ = 208$, $p = .008$, $r = .64$; Wilcoxon signed-rank tests).

Marked Cup. In the dataset including the supernumerary participants, the proportion of trials whose participant selected the marked cup as their second choice was significantly lower

than predicted by chance in the opaque condition ($M = 0.09$, $M_{dn} = 0.06$, $SD = 0.10$, $W_+ = 0$, $p < .001$, $r = -.1$; Wilcoxon signed-rank tests).

Unmarked Empty Cup. In the dataset including the supernumerary participants, the proportion of trials whose participant selected the unmarked empty cup as their second choice did not differ significantly from chance level in opaque condition ($M = 0.42$, $M_{dn} = 0.43$, $SD = 0.24$, $W_+ = 178$, $p = .097$, $r = .41$; Wilcoxon signed-rank tests).

7.4. Additional Results and Analyses: Study 3

7.4.1. Study 3's Warm-up Phase

In Study 3, the proportion of warm-up trials in which participants chose the baited container was significantly higher than that predicted by chance, (1/4) in both conditions (Ambiguous: $M = 0.85$, $M_{dn} = 1$, $SD = 0.18$, $W_+ = 153$, $p < .001$, $r = 1$; Irrelevant: $M = 0.94$, $M_{dn} = 1$, $SD = 0.11$, $W_+ = 171$, $p < .001$, $r = 1$; Wilcoxon signed-rank tests).

7.4.2. Complementary Analysis: Effect of the Shared Feature in Study 3's Ambiguous Condition

In half of the test trials of the ambiguous cue condition, the picture on the cue card had the same shape as the unique picture, and the same filling pattern as the repeated pictures; in the other half of the trials, the opposite was true. Thus, we conducted an additional analysis to assess any possible effect of the feature (i.e., shape or pattern) that the cue-card shared with the unique picture in the ambiguous cue condition. Results revealed that the feature shared by the cue-card and the unique picture (shape vs. pattern) had no effect on the proportion of trials in which the participant successfully selected the box with the unique picture ($W_+ = 33.5$, $p = .67$, $r = -.14$, Wilcoxon signed-rank test; $M = 0.35$, $M_{dn} = 0.25$, $SD = 0.23$ when the cue and unique picture had

the same pattern; $M = 0.38$, $M_{dn} = 0.25$, $SD = 0.23$ when the cue and unique picture had the same shape).