Seeing behind the surface: Communicative demonstration boosts category disambiguation in 12-month-olds

Ágnes M. Kovács[a], Ernő Téglás[a], György Gergely[a], Gergely Csibra[ab]

Affiliations:

[a] Cognitive Development Center, Central European University, Budapest, Hungary.

[b] Department of Psychological Sciences, Birkbeck, University of London, UK.

Correspondence: Ágnes M. Kovács, Cognitive Development Center, Central European University, Nádor u. 9, Budapest, 1051, Hungary. Email: kovacsag@ceu.hu

Running head: Communicative demonstration boosts category disambiguation

Research Highlights

- One-year-olds tend to represent familiar objects as exemplars of their category
- Communicative demonstration facilitates category disambiguation
- Non-verbal communication can guide infants to disregard the surface features of objects

Abstract

In their first years, infants acquire an incredible amount of information regarding the objects present in their environment. While often it is not clear what specific information should be prioritized in encoding from the many characteristics of an object, different types of object representations facilitate different types of generalizations. We tested the hypotheses that one-year-old infants distinctively represent familiar objects as exemplars of their kind, and that ostensive communication plays a role in determining kind membership for ambiguous objects. In the training phase of our experiment, infants were exposed to movies displaying an agent sorting objects from two categories (cups and plates) into two locations (left or right). Afterwards, different groups of infants saw either an ostensive or a non-ostensive demonstration performed by the agent revealing that a new object that looked like a plate can be transformed into a cup. A third group of infants experienced no demonstration regarding the new object. During test, infants were presented with the ambiguous object in the plate format, and we measured generalization by coding anticipatory looks to the plate or the cup side. While infants looked equally often towards the two sides when the demonstration was non-ostensive, and more often to the plate side when there was no demonstration, they performed more anticipatory eye movements to the cup side when the demonstration was ostensive. Thus, ostensive demonstration likely highlighted the hidden dispositional properties of the target object as kind-relevant, guiding infants' categorization of the foldable cup as a cup, despite that it looked like a plate. These results suggest that infants likely encode familiar objects as exemplars of their kind and that ostensive communication can play a crucial role in disambiguating what kind an object belongs to, even when this requires disregarding salient surface features.

Objects in our environment can be represented in various ways: as physical entities occupying a specific position in space or following a certain pathway, as a collection of perceptual features, including both intrinsic properties (such as shape, color, texture) and extrinsic or relational characteristics (such as location and orientation), or as exemplars of the object kind or category they belong to (as ‘a car’ or ‘a banana’). These representations may exist in parallel, and the context or the perceivers’ current activity may determine which of them is accessed.

From very early on, infants seem to have access to all of these types of representations. Studies on object tracking and object permanence show that young infants represent the location of objects even when they are not perceptually available (Baillargeon & DeVos, 1991; Wynn, 1992; Kaufman et al., 2003, 2005), and studies exploring perceptual categorization of serially presented stimuli show that infants represent at least some perceptual features of objects (Bushnell & Roder, 1985; Eimas & Quinn, 1994). It is more controversial whether infants rely on categorical or conceptual representation of objects as well. Some studies demonstrate, however, that towards the end of their first year, infants can represent objects as exemplars of their kinds, rather than representing them solely in terms of their perceptual features.

For example, in an object individuation paradigm, when 12-month-olds were presented with two objects of familiar categories (e.g., balls, cups, bottles), they succeeded in individuating them by their respective kinds (when the two objects belonged to two categories), but failed to do so on the basis of differences in features that did not affect their kind membership, such as size or color (Xu, Carey, & Quint, 2004). Similarly, in a paradigm introduced by Woodward (1998), one-year-olds interpreted an agent’s preference expressed in her object choice as selecting an object of a specific category rather than an object with specific features (Spaepen & Spelke, 2007). In a recent study, it has also been demonstrated that 14-month-old infants prefer to interpret a novel label as expressing the functional role that an agent plays in an interaction (e.g., chaser) rather than referring to an agent with specific features (Yin & Csibra, 2005).

While kind based, or conceptual representations may differ from perceptual representations along various aspects[1], it is often unclear how to distinguish between the two, because kind membership is normally recognized by perceptual object features (Carey, 2009). Shape, for instance, is based on a surface description of the objects we are interacting with, however, we can also integrate it in conceptual representations when developing abstract theoretical constructs (e.g., like those used in geometry, Carey, 2009; Spelke, Lee & Izard, 2010), or when we encode a shape-related feature of affordance as central for the function-based categorization of tools (Träuble & Pauen, 2007).

In many situations, there are good reasons why infants may use kind-based rather than feature-based representations. If an object of a familiar kind is occluded, representing it in terms of its kind provides a more compact representation than representing its exhaustive set of perceptual features. Object labels are more likely to refer to conceptually rather than purely featurally defined categories. People's preferences are usually about object kinds (such as types of food or kinds of toys), even if sometimes particular objects can be favorites. However, it is unclear whether infants' tendency to form kind-based representations of objects is due to a genuine bias or it is triggered by the specific context in

[1] Crucial differences between perceptual and conceptual representations may regard the content of information they encode and the nature of the inferential processes they contribute to. Perceptual representations possibly contain an inventory of the immediately experienced features of an object. The content of conceptual representations, on the other hand, depends on the domain the objects belong to. In case of human-made artifacts (such as the ones we adopted in the present study), their functions are the most central information included in such representations, though they may (and in most of the cases, should) include their stable physical features that are generalizable for the kind and can be exploited for recognizing novel kind members.

which infants are exposed to target objects. To answer this question, one could, for instance, confront infants with ambiguous situations where it remains unspecified whether it is the category or the specific features of an object that are relevant for a decision or prediction to be made.

Such a situation is created when an arbitrary rule is introduced to infants to be acquired. Infants seem to readily learn regularities based on particular perceptual features (e.g., triangle shaped objects move to a specific location, McMurray & Aslin, 2004), or on more abstract properties (e.g., auditory sequences with a specific structure predict a stimulus at a specific location, Kovács & Mehler, 2009; Hochmann et al., 2011). For example, 12-month-olds learn to anticipate the appearance of a rewarding visual stimulus that is contingently predicted by the structure of the preceding auditory sequences (Kovács & Mehler, 2009). If the rules to be used for predicting a reward rely on object categories, one can learn either that a specific category or that the features that are shared by the category members (or both) predict a particular outcome. Although both are valid descriptions of the observed events, they may differ in what kinds of generalizations they support for objects whose category membership is ambiguous.

How would infants find the correct basis for generalization when there is more than one option available? Only a few studies attempted to expose infants to multiple regularities simultaneously. In studies where two generalizations were possible, infants selected as the basis for generalization the one that was statistically more likely, or more salient given a specific set of stimuli (Gerken, 2006; Kovács & Mehler, 2009; Kovács, 2014). However, such learning must rely on a considerable amount of exposure and on extensive sampling of the available information. Infants are highly sensitive to the co-occurring patterns present in their linguistic (Saffran et al., 1996; Gomez & Gerken, 1999), visual (Fiser & Aslin, 2002; Kirkham, Slemmer & Johnson, 2002) or social input (Bahrick & Watson, 1985). Besides detecting statistical relations between co-occurring items or events, they can also perform generalizations and inferences regarding items never encountered before. Still, the stimulus itself often may not carry enough statistical information to trigger the most appropriate inference, and in such cases infants would benefit from being able to exploit other types of information as well.

Various proposals have been put forward regarding the factors that may trigger conceptual representations of objects in infants, focusing on the role of labels, function demonstration, and ostensive communication. For example, it was suggested that linguistic input accompanying object presentations guides early learning about object kinds. Adult communication containing verbal labels seems to provide strong cues for kind-based inferences for young children. If an adult labels an animal for the child, 2-year-olds perform category-based inferences guided by essential behavioral properties (e.g., what an animal feeds on), while disregarding perceptual features (Gelman & Coley, 1990). Importantly, children in this study ignored the label when it named a transient property rather than a stable category. Furthermore, even infants as young as 9 months seem to capitalize on principles inherent to linguistic labeling. In an object individuation paradigm they considered labels as referring to object kinds, thus labeling two objects with different names induced the expectation that two objects should be present (Xu, 2002, 2005).

Linguistic labels, however, might not be unique in guiding infants' expectations, as they also seem to treat artifact function as a kind determining property. Causal understanding of artifacts provides infants with sufficient information to guide differential function-specific expectations. It has been shown that 12-month-olds categorize a set of unfamiliar objects according to a perceptually non-salient part while ignoring more salient parts, but only do so following a demonstration of the functional relevance of this part (Träuble & Pauen, 2007). Recent studies have found that nonverbal ostensive function demonstration, even in the absence of causal information, also facilitates infants' encoding of artifact functions (Futo et al., 2010), and shields learning about the functions of tool kinds from counter-evidence (Hernik & Csibra, 2015).

Our study tested the hypotheses according to which one-year-old infants tend to encode familiar objects as exemplars of their kind even when lower-level featural encoding would also be possible and sufficient, and that ostensive communication plays a role in disambiguating the specific kind that an ambiguous object belongs to. Specifically, we investigated whether observing an ostensive demonstration revealing that a plate-like object is in fact a foldable cup would allow infants to categorize it as a cup even when they see it again in the plate format. If so, ostensive demonstration would play a kind-specifying role, fixing the non-obvious but kind-relevant properties of the object that do not have to be directly perceived, and driving disambiguation similarly to the verbal assertion "This may look like a plate, but it is in fact a cup." We contrasted this condition with a non-ostensive demonstration and with a third condition that involved no demonstration at all.

We performed an eye tracking study comprised of a training phase, a demonstration phase and a test phase. We first trained infants in a categorization task: they were exposed to movies displaying an agent consistently sorting objects from two categories (cups and plates) into two locations (left or right), where a rewarding puppet also appeared. Then a demonstration phase followed, in which the agent performed an ostensive or non-ostensive demonstration to reveal that a newly introduced object could be transformed from a plate format to a cup format. In a third condition infants were exposed to no demonstration at all regarding this new object. During the test phase of all three conditions, infants were presented with the target object in the plate format, and we assessed whether they anticipated the puppet's appearance on the cup- or the plate-side.

If infants learn the regularity between the objects and the location of the reward at the level of the visual features of the objects (e.g., objects that are flat such as plates, go to one location, and objects that are elongated, such as cups, go to the other location), infants should anticipate to the plate-side in all conditions, since the foldable cup appears in the plate-like surface format in the test. However, if they learn the regularity at the level of the object categories (cups and plates), and ostensive communication during the demonstration is interpreted as revealing the object kind that the referent belongs to, infants should categorize the ambiguous object as a cup and make anticipatory looks to the cup side. This categorization, and the ensuing anticipation, is not expected in the other two conditions. Thus, if ostensive demonstration, as hypothesized, highlights hidden dispositional properties of the plate-like object that are kind-relevant, it should guide infants' categorization of the foldable cup as a cup, despite that it sometimes looks like a plate (Csibra & Gergely, 2009).

Methods

Participants

Forty-eight 12-month-old infants participated in the study. Participants were randomly assigned to one of the following three conditions (16 infants per condition): Ostensive Demonstration (OD, mean age 12.21, age-range 12.05 -13.02, 6 boys); Non-ostensive Demonstration (NOD, mean age 12.17, age-range 12.03 – 12.31, 6 boys); and No Demonstration (ND, mean age of 12.21, age-range 12.06 -12.29, 9 boys). All infants were full term and were recruited on the basis of local birth records. They received a small toy gift for their participation at the end of the experiment. The parents gave informed consent for the participation in the experiment. Further 8 infants were not included in the data analysis because of inattention or fussiness (OD: 3; NOD: 3; ND: 2), 4 infants because of failure of recording eye gaze data (OD: 1; NOD: 2; ND: 1), and 3 infants because of other technical problems (OD: 2; ND: 1).

Apparatus

The gaze data was collected at 60 Hz by a Tobii T60 Eye Tracker (Stockholm, Sweden). The stimuli were presented on a 17-inch monitor that contained an inbuilt eye tracker. The stimulus presentation was performed with PsyScope X (http://psy.ck.sissa.it; Cohen, MacWhinney, Flatt, & Provost, 1993) on a Mac Pro. Psyscope X sent event log information to the gaze data file (recorded with ClearView, running on a separate computer) marking the event sequences in the trials.

Stimuli and Procedure

Infants sat on their caretakers' lap approximately 60 cm from the screen in a dimly lit sound-proof room. Caretakers wore opaque sunglasses during the experiment and were instructed not to interact with the infants. The experimenter monitored infants' behavior via a video camera. The experiment consisted of training trials, demonstration trials, and test trials. The three experimental groups had identical training and test trials, and differed only in the demonstration trials. The experiment was preceded by the calibration of the eye tracker applying a five-point calibration routine.

The experiment had the following structure: there were two main blocks each containing 10 training trials, 2 demonstration trials (in the OD and NOD conditions only), 2 further training trials, followed by 4 test trials. All trials were preceded by a central attention getter used to attract the infant's attention to the middle of the screen.

In the training trials, infants were shown that cups and plates were consistently sorted to a particular side (left or right, counterbalanced across infants) of the screen. They watched short movies (7 s) displaying a person positioned centrally at a table, 2 red occluders on the two sides of the table, and a target object in the middle of the table. The object was an exemplar of two object categories (plates and cups) of different shapes and colors. The person lifted the object from the table up to chest level and stopped in this position (for 1 second). Then, a puppet appeared from behind one of the two occluders at the left or the right side accompanied by a tinkling sound (see Supplementary Material Movie S1). Afterwards the person placed the object behind the occluder from where the puppet appeared (see Figure 1). The person kept her eyes on the object and never looked to the infant during the training trials. For half of the infants cups were consistently paired with a puppet on the left side and plates with a puppet on the right side, and for the other half of the infants the opposite pairing was used. We used 4 cups and 4 plates randomly paired with 2 puppets (a rooster and an elephant) presented in a pseudo-random order (no movie was repeated immediately and a maximum of 2 consecutive trials were presented from the same category).

After 10 training trials, the demonstration trials followed (except in the No Demonstration condition). Infants watched 9-second-long movies, in which the same person appeared in the same setting as in the training trials, with the exception that now there was a new plate-like object (target object) displayed centrally. During the Ostensive Demonstration condition, the person looked towards the viewer and addressed the child in infant-directed intonation, saying “Hi baby, hi! Look!” She then lifted the plate-like object and demonstrated that it was in fact a foldable cup by transforming it into its cup-shape, looked at the child again, and said “Aha” in an intonation suggesting satisfaction. The movies in the Non-ostensive Demonstration condition had the same timing, and the same manipulation was performed on the target object, except that the agent never looked at the infant and used adult-directed intonation, saying “Let’s see what’s this!” and “Aha” (with an intonation suggesting recognition) after the manipulation. In both conditions, the demonstration trial finished at this point, i.e., no puppet appeared and the target object was not moved to either side (see Figure 1).

The demonstration trials were followed by two further training trials (a cup trial and a plate trial) in a counterbalanced order to remind infants of the object-side pairing. In the No Demonstration condition, these trials came right after the previous training trials, making the training phase 12-trial long.

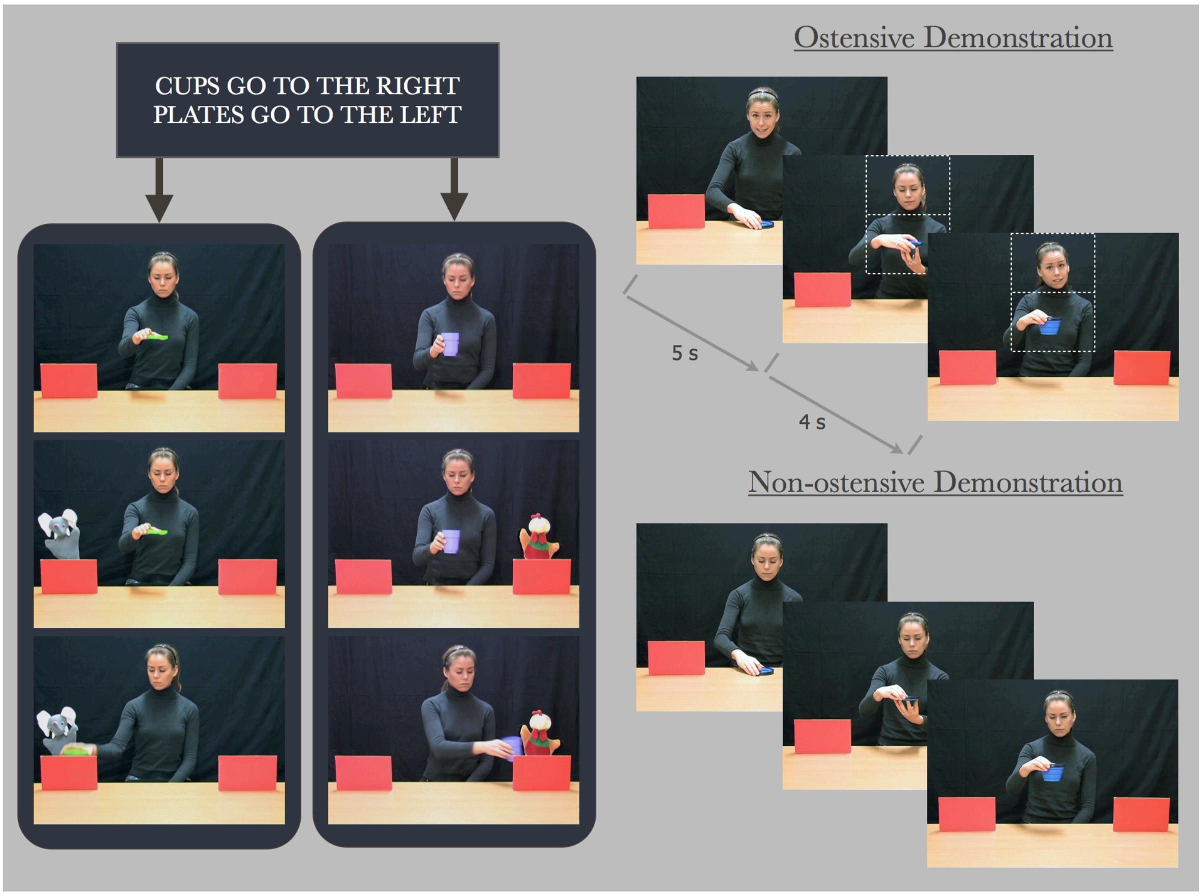

Figure 1. Snapshots of the videos presented in the training and demonstration phase. Left panel: Training trials illustrating the rules of sorting for the two object categories. Right panel, upper part: Ostensive Demonstration of object transformation (the actor looks to the infant and addresses her in infant-directed speech then performs the transformation). Dashed rectangles depict the regions on which the AOI analysis was performed (face AOI vs. object AOI) in the demonstration phase. Right panel, lower part: Non-ostensive Demonstration of object transformation (the actor does not look to the infant and uses adult-directed speech). The timing and the AOIs were the same in the Ostensive Demonstration and the Non-ostensive Demonstration conditions

The four test trials, which were identical in all the conditions, presented infants with 3-

second-long movies. These movies were identical to the training movies except that (1) the person now lifted the target object displayed in the plate format and held it at chest level, at which point (2) the tinkling sound associated with the puppets were played to prompt the appearance of a puppet, which, however, never appeared. At this moment, the screen froze for 2 seconds. Note that sorting was not performed during the test trials, and the target object was not placed behind the occluders. This procedure allowed us to measure infants anticipatory looks towards one or the other occluder, reflecting their expectation about the location where the puppet would appear during the two seconds after hearing the tinkling sound that had been associated with the appearance of the puppets. Thus, in the test trials we assessed whether infants exhibited categorization of the target object seen in the plate format as a plate or as a cup by looking towards the side they had associated with these two categories during the training phase.

Results

We assessed to which category infants thought the target object belonged by measuring their first look after hearing the tinkling sound. If infants categorized the target object as a plate, or learnt the arbitrary rule for the shape of the objects rather than for their categories, they should perform anticipatory looks to the side where the puppet used to appear for plates. However, if they categorized it as belonging to the category of cups, they should look to where the puppet used to appear for cups. The screen was divided into 3 equal parts on the horizontal axis: left, middle, and right. Looks to the middle part (corresponding to location of the person) were excluded. We coded infants' first looks to the left or right side of the screen during the 2-s time window after the tinkling sound was played. We calculated difference scores across the 8 test trials [(number of trials with anticipation to the cup side minus number of trials with anticipation to the plate side) divided by (number of trials with anticipation to the cup or plate side summed)] for each condition. We tested these scores against the chance level of 0 (no preferential anticipation), and compared them across conditions. Positive values reflect that infants looked more often to the cup side, while negative values indicate that they looked more often to the plate side in the test.

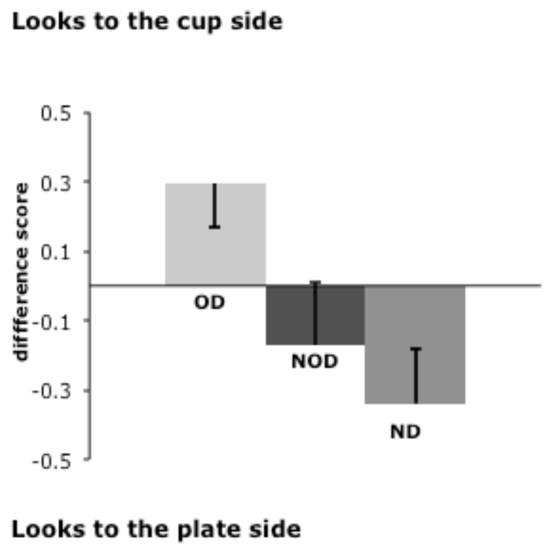

Figure 2. Anticipatory looks (first looks) to the cup and plate sides as measured in the 3 experimental conditions: Ostensive Demonstration-OD, Non-ostensive Demonstration-NOD, and No Demonstration-ND

Infants in the Ostensive Demonstration condition looked more often to the cup side than to the plate side in the test trials [mean difference score 0.29, tested against zero: $t(15) = 2.32$, $p = .03$; Fig. 2]. However, infants in the Non-ostensive Demonstration condition showed anticipatory looks that were not different from chance [mean difference score -0.17, tested against zero: $t(15) = 0.95$, $p = .35$]. When comparing the Ostensive Demonstration condition with the Non-ostensive Demonstration condition, infants displayed a significantly different looking pattern [$t(30) = 2.12$, $p = .04$] suggesting more anticipation towards the cup side in the OD condition (Figure 2).

Comparing anticipatory looks in the No Demonstration condition to chance level, we found that infants looked significantly more often to the plate side [mean difference score -0.34, $t(15) = 2.19$, $p = .04$]. Infants showed different anticipatory looks in the Ostensive Demonstration and No Demonstration conditions, $t(30) = 3.16$, $p = .003$, however, looking patterns were not significantly different between the Non-ostensive Demonstration and No Demonstration conditions, $t(30) = 0.71$, $p = .48$.[2]

The proportion of total anticipatory looks to the cup or plate side in the test was highly similar in the three conditions (.42 in the Ostensive Demonstration condition, .41 in the Non-ostensive Demonstration condition, and .35 in the No Demonstration condition, these values are not significantly different from each other). Two infants in the OD condition, 1 infant in the NOD condition and 2 infants in the ND condition did not perform any anticipatory looks (nor correct, neither incorrect) in the test trials. An analysis where we excluded these infants from the comparisons yielded practically identical results. In the trials where the infants did not perform anticipatory looks, they tended to fixate to the middle of the screen where the agent was displayed. Thus, the differences we found across the three conditions could not have been due to different amount of motivation to anticipate the puppets in the three conditions. Similar proportion of anticipations is not unusual in eye tracking research with infants; previous studies reported anticipations from .25 to .45 (Canfield et al., 1997; Johnson, Amso, & Slemmer, 2003).

We then analyzed infants' looking behavior in the demonstration phase of the OD and NOD conditions, to investigate whether infants in the NOD condition were at chance during

[2] Difference scores computed on overall looking duration (rather than first looks) to the two target areas yield a similar pattern (OD condition: M = 0.23, SD = 0.65; NOD condition: M = -0.12, SD = 0.77; ND condition: M = -0.18, SD = 0.78, differences tending towards, but not reaching significance). Thus, similarly to other studies, first looks reflected between-condition differences more reliably than looking duration (Hochmann et al., 2011; Senju & Csibra, 2008; Rubio-Fernandez & Glucksberg, 2012).

the test because they looked less to the object than infants in the OD condition, and hence failed to encode its transformation. Thus, we analyzed the distribution of infant's attention between the object and the face areas of interest (AOIs, 240 x 340 pixels each) in the demonstration phase. The time window of this analysis included the entire length of object transformation: it started when the agent lifted the plate-like object and lasted until the end of this event (period of 4 s, see the right panels of Figure 1). We first calculated the proportion of time infants looked to the object AOI compared to the time they spent in looking to the object AOI plus the face AOI. We found that, interestingly, infants in the NOD condition looked significantly more to the object AOI (average proportion 0.71) than infants in the OD condition (average proportion 0.44) ($t(30) = 4.73$, $p < .0001$). Next, we computed how much infants in the two conditions looked to the object AOI on average. Infants in the NOD condition tended to look longer to the object AOI (mean = 1832 ms) as compared to infants in the OD condition (mean = 1416 ms), ($t(30) = 1.69$ $p = .10$). These results exclude the possibility that infants were at chance in the test of the NOD condition because they failed to follow the transformation, as they spent more time looking to the object AOI in this condition than infants in the OD condition. These findings also suggest that what infants encode from an event may depend more on the context and the triggered learning mechanism, rather than on how much time they spend on inspecting various parts of the scene.

Finally, we addressed the question whether infants' chance level anticipation in the test phase at group level in the NOD condition was due to general uncertainty of categorization on the individual level, or to an even distribution of cup-categorizers and plate-categorizers. Unfortunately, the small sample size did not allow us to test the uni- or bimodality (respectively corresponding to the two options above) of the distribution of difference scores. Across the test trials, some infants (n = 4) performed more anticipations to the cup side, some (n = 8) performed more anticipations to the plate side, and others (n = 4) preferred the two sides equally. To better understand these data, we correlated the time infants in the NOD condition looked to the object AOI in the demonstration phase to their difference score in test. Infants' looking pattern to the object transformation in the demonstration phase was not related to their performance in test (r = -0.03).

Discussion

While the test phase was identical in the three conditions, i.e., infants were presented with the ambiguous object in the *plate* format, infants seemed to make different generalizations in the three conditions. In the No Demonstration condition they anticipated the puppet on the plate side, suggesting that they either recognized the target object as belonging to the category of plates or they had learnt the side rule to the shape of the objects. This result demonstrates that the training provided for infants was sufficient to learn the rule linking the two categories or the two shapes to the corresponding side of the screen.

In the Non-ostensive Demonstration condition infants, as a group, looked equally often towards the two sides. Although this pattern was not statistically significantly different from that of the ND condition, it did not differ from chance level either. This suggests that at least some of the infants in this condition may have recognized the object as the same one as they had seen in the demonstration phase, and this made them hesitant to choose which category it belonged to.

In contrast, infants in the Ostensive Demonstration condition performed more anticipatory eye movements in the test to the *cup* than to the *plate* side. Thus, they recognized the object as the same one they had seen during the demonstration trials, and inferred that it belonged to the category to which the object was transformed during the ostensive demonstration, rather than to the one it belonged to on the basis of its current appearance. Generalizing a specific rule (e.g., “cups go to right”) to a new member of the category despite it displaying contradicting visual features points to the possibility that in the training phase (which was entirely non-ostensive), infants had already encoded the cup-related information via kind-based representations. Only kind-based learning, but not feature-based learning could allow infants to generalize the “cups go to right” rule to a member of the cup category that looks like a plate but can be transformed into a cup. The fact that one-year-old infants are familiar with these kinds of objects, and most of them have even acquired the word for at least one of them (‘cup’ in English: Gliga & Csibra, 2009, in Hungarian: Parise & Csibra, 2012) might have contributed to this result.

In the two conditions involving a demonstration regarding the target object a non-obvious property of the object was revealed (a plate transformed into a cup). However, when seeing the object again in the plate format, only the ostensive demonstration condition led infants to expect that this object would go to the cup side. The most likely explanation we see for why infants would go with this hidden feature is that ostensive communication marked the cup-shape outcome of the transformation as a kind relevant one, crucial for categorization.

Our results suggest that ostensive communication can play a role in disambiguating what kind an object belongs to from early infancy, even if this requires disregarding its surface features. This conclusion is in line with findings with preschoolers, who go beyond salient perceptual features (color or shape) when sorting objects only if object function is ostensively demonstrated to them (Butler & Markman, 2014). It seems that ostensive demonstration serves a similar function already in infancy, at least with familiar object kinds. We started this study from the general hypothesis that infants would be biased to form conceptual representations for ostensively referred objects. While this study cannot prove this hypothesis directly, our results successfully excluded the alternative hypothesis, according to which infants form solely perceptual representations (and perhaps perceptual categories) from the immediately available features of objects regardless of the context. Our findings suggest that the object representations that infants developed in the OD condition were not merely perceptual, as they seemed to disregard the actual surface characteristics of the objects they observed during the test phase.

How does ostensive demonstration exert its effect? After all, infants in the NOD condition had the same amount of evidence of the ambiguous nature of the foldable cup as infants in the OD condition. However, if ostensive communication makes infants expect that the content of the communication will reveal kind-relevant information for them about the referent (Csibra & Gergely, 2009; Csibra & Shamsudheen, 2015), they could interpret the final state of the demonstration as the predicate of the message (possibly making the demonstration the non-linguistic equivalent of the sentence “This is [in fact] a cup”). Seemingly, ostensive communication by itself can serve as a kind-defining assertion. In contrast, the mere observation of the changing object in the NOD condition would not tell infants whether the object is a ‘foldable cup’ or an ‘extendable plate’ (if either). Thus, while it provides information about the two states that the object can take, it does not specify which state is the one that represents the category the object belongs to.

One might wonder how such a possible ‘dual-identity object representation’ may differ from the unambiguous kind representation infants may form in the OD condition. While even the ‘dual-identity representations’ would be likely conceptual, it is unclear whether infants can form such representations. Data from other domains suggest that multiple descriptions are usually avoided. Infants and children, for instance, avoid multiple labels for a single object (Markman & Wachtel, 1988). We believe that infants are more likely use their pre-existing concepts or categories (such as cup, or plate) in forming arbitrary associations and generalizations, than concepts created on the fly (e.g., ‘tall objects’ and ‘flat objects' in the familiarization, and 'transformable plate' or 'dual identity object' in the test). In either case, the finding that infants categorized the plate-like target object to the cup category in the OD condition suggest that they have formed a representation that went beyond the actual perceptual features of the object.

Young children may often face situations of ambiguity where multiple representations, hence multiple types of generalizations, are possible. The question is whether (and how) the developing cognitive system dedicated to information acquisition gives priority to one or another kind of representation. Here we have shown that the representation of objects in terms of their kinds rather than in terms of their actual perceptual features enjoys a priority at the end of the first year of life. In addition, we found evidence that ostensive communication can play a role in kind disambiguation. Linguistic labels have been known to carry kind information for infants (Xu, 2002; Waxman & Markow, 1995), but it seems that non-verbal aspects of communication can also accomplish the same function. Other types of social cues might as well guide learning about 'what an object is for' in the context of cultural conventions or social groups (Diesendruck, & Markson, 2011; Buttelmann et al., 2013). Whether the effects we observe in the present study are restricted to categories that are well-established or even lexicalized in the infant mind (such as 'cups'), or it extends to novel, perceptually based categories, is a question for further research.

Acknowledgements:

We thank the families who participated in this study, and Mikołaj Hernik for comments on a previous version of the manuscript. This research was supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Advanced Investigator Grant (249519-OSTREFCOM) to G.C., and partly by an ERC starting grant to Á.M.K. (284236- REPCOLLAB) and an ERC Synergy grant (609819 – SOMICS) to G.G.

References:

Bahrick, L. R., & Watson, J. S. (1985). Detection of intermodal proprioceptive-visual contingency as a potential basis of self-perception in infancy. *Developmental Psychology*, *21*, 963-973.

Baillargeon, R., & DeVos, J. (1991). Object permanence in 3.5- and 4.5-month-old infants: Further evidence. *Child Development, 62,* 1227–1246.

Bushnell, E. W., & Roder, B. (1985). Recognition of color-form compounds by 4-month-old infants. *Infant Behavior and Development, 8,* 255–268.

Butler L.P, & Markman E.M. (2014). Preschoolers use pedagogical cues to guide radical reorganization of category knowledge. *Cognition, 130*, 116–27.

Buttelmann, D., Zmyj, N., Daum, M. M., & Carpenter, M. (2013). Selective imitation of in-group over out-group members in 14-month-old infants. *Child Development*, *84*, 422-428.

Canfield, R. L., Smith, E. G., Brezsnyak, M. P., & Snow, K. L. (1997). Information processing through the first year of life. *Monographs of the Society for Research in Child Development,* 62.

Carey, S. (2009). The origin of concepts. New York: Oxford University Press.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, & Computers, 25,* 257–271.

Csibra G., & Shamsudheen R. (2015). Non-verbal generics: Children’s interpretation of object-mediated communication. *Annual Review of Psychology,* 68:689-710.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*, 148–153.

Diesendruck, G., & Markson, L. (2011). Children's assumption of the conventionality of culture. *Child Development Perspectives*, 5, 189-195.

Eimas, P., & Quinn, P. (1994). Studies on the formation of perceptually based basic- level categories in young infants. *Child Development, 65*, 903–917.

Fiser, J. and Aslin, R.N. (2002) Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, U. S. A. 99*, 15822–15826

Futó, J., Téglás, E., Csibra, G., & Gergely, G. (2010). Communicative Function Demonstration induces kind-based artifact representation in preverbal infants. *Cognition*, *117*, 1–8.

Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology*, *26*, 796–804.

Gerken, L.A. (2006). Decisions, decisions: infant language learning when multiple generalizations are possible. *Cognition*, 98, B67–B74.

Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science, 20*, 347-353.

Gómez, R.L., & Gerken, L.A. (1999). Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. *Cognition*, 70, 109–135.

Hernik M. & Csibra G. (2015). Infants learn enduring functions of novel tools from action demonstrations. *Journal of Experimental Child Psychology,* 130, 176-192.

Hochmann J.R., Benavides-Varela S., Nespor M., & Mehler J (2011). Consonants and vowels: Different roles in early language acquisition. *Developmental Science*, 14, 1445-58. doi: 10.1111/j.1467-7687.2011.01089.x

Johnson, S., Amso, D., & Slemmer ,J. (2003) .Development of object concepts in infancy, Evidence for early learning in an eye tracking paradigm. *Proceedings of the National Academy of Sciences ,USA, 100*, 10568–10573.

Kaufman, J., Csibra, G., Johnson, M.H. Oscillatory activity in the infant brain reflects object maintenance. . *Proceedings of the National Academy of Sciences ,USA 102*, 15271-15274 (2005).

Kaufman, J., Csibra, G., Johnson, M.H. (2003). Representing occluded objects in the human infant brain. *Proc. R. Soc. Lond. B Biol. Sci.*, *270***,** (Suppl 2), S140-S143

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy. *Cognition*, *83*, B35–B42.

Kovács, Á. M. (2014). Extracting regularities from noise: Do infants encode patterns based on same and different relations? *Language Learning,* 64, 65-85, DOI: 10.1111/lang.12056

Kovács, A. M., & Mehler, J. (2009). Flexible learning of multiple speech structures in bilingual infants. *Science*, *325*, 611–612.

Markman, E. M. & Wachtel, G. A. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20,* 120-157.

McMurray, B., & Aslin, R.N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy*, *6*, 203–229.

Parise E., & Csibra G. (2012). Electrophysiological evidence for the understanding of maternal speech by 9-monthold infants. *Psychological Science, 23*, 728–33.

Rubio-Fernández, P., & Glucksberg, S. (2012). Reasoning about other people's beliefs: Bilinguals have an advantage. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *38*, 211–217.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month old infants. *Science*, *274*, 1926–1928.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology, 18*, 668-671.

Spaepen, E. & Spelke, E. (2007). Will any doll do? 12-month-olds' reasoning about goal objects. *Cognitive Psychology, 54,* 133-154. DOI: 10.1016/j.cogpsych.2006.06.001

Spelke, E., Lee, A. A., & Izard, V. (2010). Beyond core knowledge: Natural geometry. *Cognitive Science, 34*(5), 863-884. DOI: 10.1111/j.1551-6709.2010.01110.x

Träuble, B. & Pauen S. (2007). The role of functional information for infant categorization. *Cognition,* 105, 362–79.

Waxman, S.R., & Markow, D.B. (1995). Words as invitations to form categories: evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*, 257–302.

Woodward, A.L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1–34.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358***,** 749

Xu, F. (2002). The role of language in acquiring kind concepts in infancy. *Cognition*, *85*, 223-250.

Xu, F. (2005). Categories, kinds, and object individuation in infancy. In *Building Object Categories in Developmental Time*, ed. L Gershkoff-Stowe, D Rakison, pp. 63–89. Mahwah, NJ: Erlbaum

Xu, F., Carey, S., & Quint, N. (2004). The emergence of kind-based object individuation in infancy. *Cognitive Psychology, 49,* 155–190.

Yin, J. & Csibra, G. (2015). Concept-based word learning in human infants. Psychological Science, 26, 1316-1324. http://dx.doi.org/10.1177/0956797615588753