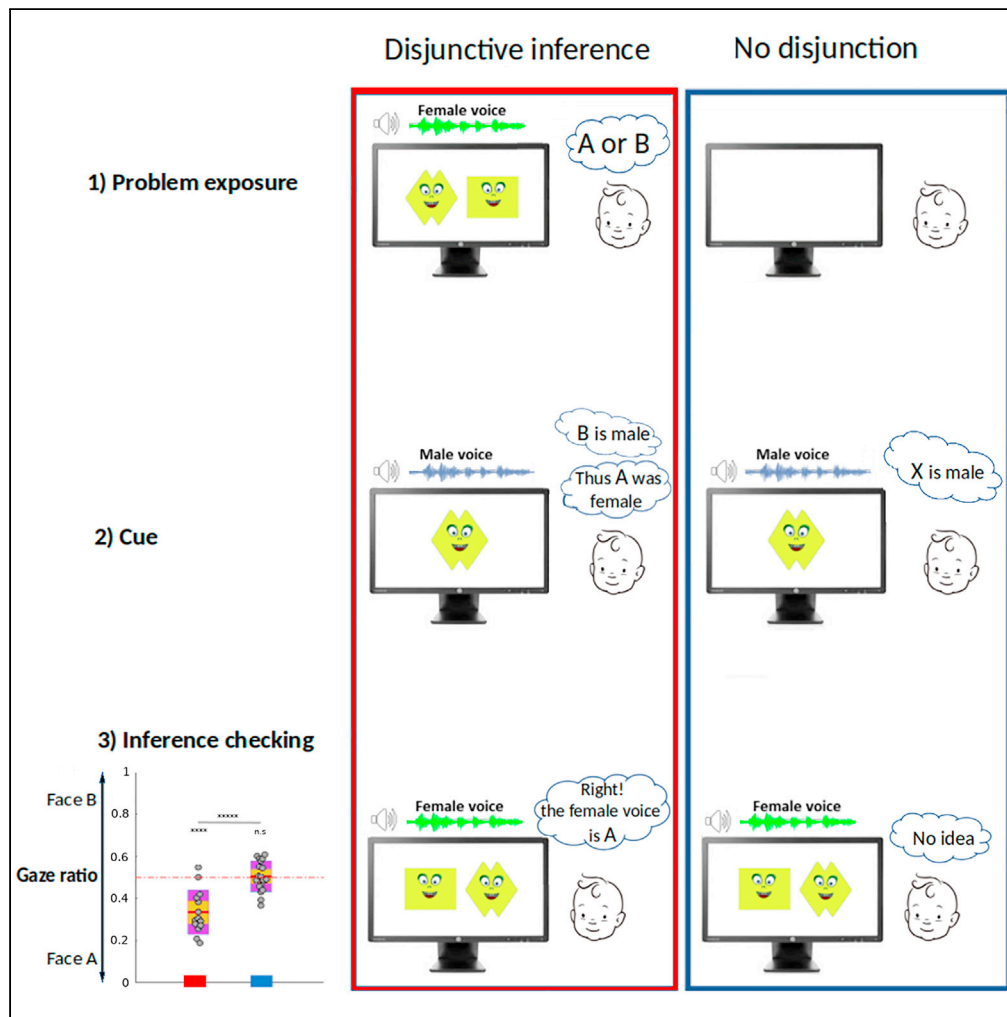


Article

Disjunctive inference in preverbal infants



Milad Ekramnia,
Jacques Mehler,
Ghislaine
Dehaene-
Lambertz

milad.ekramnia@cea.fr

Highlights

10-month-old infants have no logical operators in their lexicon

Nevertheless, they can use logical deduction in case of an ambiguous situation

They correctly deduce which faces and voices are paired through disjunctive inference

Infants' performance in this task can be followed by measuring their pupil dilation



Article

Disjunctive inference in preverbal infants

Milad Ekramnia,^{1,2,3,*} Jacques Mehler,^{2,4} and Ghislaine Dehaene-Lambertz¹

SUMMARY

Can preverbal infants utilize logical reasoning such as disjunctive inference? This logical operation requires keeping two alternatives open (A or B), until one of them is eliminated (if not A), allowing the inference: B is true. We presented to 10-month-old infants an ambiguous situation in which a female voice was paired with two faces. Subsequently, one of the two faces was presented with the voice of a male. We measured infants' preference for the correct face when both faces and the initial voice were presented again. Infant pupillary response was measured and utilized as an indicator of cognitive load at the critical moment of disjunctive inference. We controlled for other possible explanations in three additional experiments. Our results show that 10-month-olds can correctly deploy disjunction and negation to disambiguate scenes, suggesting that disjunctive inference does not rely on linguistic constructs.

INTRODUCTION

The world is complex but not random. How can human children make sense of the myriad of interactions between events, some causal, some not, to which they are exposed? Understanding these events would be greatly simplified if they had the capacity to predict the consequences of certain events based on previous knowledge (i.e., "If my sister likes chocolate and my brother likes candy, and if there is no more chocolates in the sweet box but still candies, my sister likely visited the box already."). Using this form of logical deduction, it is thus not necessary to see the event to deduce what happened and propose an explanatory chain for the current event (i.e., a box with only candies). But to make this inference, it is necessary to have access to logical operators such as "if," "and," "or," "not," "then," etc. ... to turn the knowledge of past events into steps of a logical reasoning (i.e., if A or B [step 1], not A [step 2], then B [step 3]).

It is currently unknown at what age children are able to deploy such logical operations. Many authors, following Piaget have considered that the mastery of language was a necessary step for such reasoning (Inhelder and Piaget, 1958). The rich symbolic and hierarchical structure of language might support the discovery of abstract relations between events. Furthermore, possessing a specific lexicon improves performance in the related tasks. Many experiments have emphasized how words may help infants to better understand and memorize a situation. For example, naming different exemplars of a category by the same word helps infants to represent this category (Fulkerson and Waxman, 2007). Thus, the production and comprehension of logical words is certainly a milestone in the development of reasoning abilities. Furthermore, several experiments have shown reasoning limitations in young children. For example, when an experimenter drops an object into an inverted Y-shaped tube, children under 4 years of age fail to reach out with both arms simultaneously to retrieve the object, supporting the hypothesis that younger children cannot keep in mind two hypotheses in order to select the correct action (i.e., the object is coming out of the left OR right arm of the tube). This and other experiments (Mody and Carey, 2016) with similar results led Leahy and Carey (2020) to conclude that the mastery of disjunctive inference is not completed before 4 years of age, owing to difficulties children experience in keeping open several possible explanations, and thus when there is an ambiguity, young children randomly choose one possibility and stick to it until invalidating evidence forces them to change and upgrade their mental model. However, limitations in working memory, executive function, control of bimanual movement, or poor understanding of the situation might also contribute to past failures of young children to exercise disjunctive inference. Therefore, a major challenge is to design appropriate experimental paradigms for young children and infants, in which these limitations would have a minimal impact on measures of reasoning abilities.

The debate has recently been reignited by the recent experimental observations suggesting the demonstration of abstract reasoning skills long before the development of language skills. For instance, 5-month-olds were able

¹Cognitive Neuroimaging Unit, CNRS ERL 9003, INSERM U992, CEA, Université Paris-Saclay, NeuroSpin center, 91191 Gif/Yvette, France

²Language, Cognition, and Development Laboratory, Scuola Internazionale di Studi Avanzati, Trieste, Italy

³Lead contact

⁴Prof. Jacques Mehler unfortunately passed away in 2020.

*Correspondence:

milad.ekramnia@cea.fr

<https://doi.org/10.1016/j.isci.2021.103203>



to notice in a few trials that never repeated tri-syllabic non-words that shared a common repetition structure (AAB, ABA, or ABB) were associated with a specific image depending on their structure (e.g., ABA triplets were followed by the image of a red fish, ABB by a lion, and AAB randomly by either image). Using EEG, it was possible to follow the chain of processing steps: rule extraction, expectation (or not) of a specific image, sensory priming for the congruent image, and late surprise if the expectation was not fulfilled, revealing that preverbal infants were able to keep three rules in mind. By itself, this result was already an achievement given the amount of training required in other species for this type of abstract learning (Ghirlanda et al., 2017). Moreover, infants immediately generalized to reversed pairs. If, for example, infants had learned that ABB triplets were followed by a red fish, when shown a red fish, they expected the following triplet to obey the ABB rule, without needing additional training to learn the reverse association, as if the image was representing the rule. This immediate bidirectional relation, not predicted by associative learning nor observed in animals (Medam et al., 2016), was proposed by the authors as a possible marker of symbolic representations in pre-verbal infants (Kabdoune and Dehaene-Lambertz, 2019). Moreover, Cesana-Arlotti et al. (2018) examined whether 12-month-olds were capable of more complex reasoning involving a logical disjunction between two alternatives. The authors implemented the three steps of a disjunctive inference situation as follows: two objects *A* and *B* were placed behind an occluder. Subsequently, one of the two was scooped from behind the occluder and put in a cup without the infant being able to see which one, potentially presenting an ambiguity for the observer (phase 1: problem exposure. *A or B is in the cup*). Therefore, if *B* appears from behind the occluder (phase 2, cued phase), it implies by inference that it should be *A* in the cup. Twelve-month-olds were indeed surprised (i.e., showed longer looking time) when the final outcome (*A or B in the cup*) was inconsistent with the prediction (phase 3, confirmation or violation of the inference). The authors interpreted the results as evidence for disjunctive inference. Similarly, 14-month-olds were able to deduce where the desired object was located among two hidden ones, after the position of the other was revealed (Cesana-Arlotti, et al., 2020). However, given the failures in toddlers, an interpretation that involves disjunctive inference was questioned and two alternative explanations have been proposed. Explanation 1: Infants were just using an object tracking mechanism at phase 3 without the need to consider the initial two possibilities in phase 1. Because an object is attached to a single location (Xu, et al., 2004), if object *A* comes out of the occluder in phase 2, it could not be in another spatial location (i.e., inside the cup, in phase 3). This explains the infant's surprise if the object in the cup and behind the occluder was the same object (Jasbi, et al., 2019). Explanation 2: At phase 1, infants only assumed one hypothesis, for example "*A is in the cup*" (*B is not represented*). If *B* appears in phase 2 from behind the occluder at another location than *A*, there is no conflict and the infant's mental model is now "*B is behind the occluder*" (*A is no more represented*). In phase 3, *A* is revealed in the cup, at another location than *B*, thus again no conflict and no surprise ("*A is in the cup*"). The second possibility is that after phase 1 where *A* was assumed in the cup, it is *A* that appears from behind the occluder in phase 2. In this case, the infant is surprised because the same object cannot be at two positions and so they cannot upgrade their mental representation with *A* being behind the occluder. In phase 3, *A* is revealed in the cup, the baby is again surprised for the same reason as in phase 2. In this scenario, coherent with Cesana-Arlotti et al.'s results, toddlers have only a single object representation at each phase. Thus, there is no disjunctive inference, which needs a minimum of two hypotheses. The two explanations differ slightly from each other, but they are crucially opposed to the disjunctive inference hypothesis, which requires the infant to represent two alternative situations simultaneously and to reason on internal representations in order to deduce an outcome as information comes in. In this case, phase 3 is only the confirmation of the elimination of one of the possibilities at phase 2.

To contribute to this debate, we present here data obtained in 10-month-old preverbal infants using oculomotor measures. Our goal was to minimize scene complexity and working memory load while at the same time teasing apart the various stages of infant reasoning. We also avoided object displacement to discard explanations based on object tracking mechanisms (Jasbi et al.'s critique) and based our paradigm on infants' social abilities. From 3 months on, infants can associate a voice to a face (Jordan and Brannon, 2006; Bahrick, et al., 2005; Brookes et al., 2001), and they apply a mutual exclusivity principle robustly after 8 months of age: i.e., when they already know a face-voice pair, they associate a new voice to a new face (Orena and Werker, 2021). Therefore, we hypothesized that it might be possible to create an ambiguous situation if two faces were presented while a voice was speaking. Our goal was to observe how infants resolve the initial ambiguity (i.e., to which face the voice belongs), by manipulating the subsequent evidence about the associations between voices and faces.

Specifically, the three logical steps are presented in Figure 1: Phase 1: two cartoon faces were presented accompanied by a female voice (*A or B*). In the example of Figure 1, the voice belongs to either the

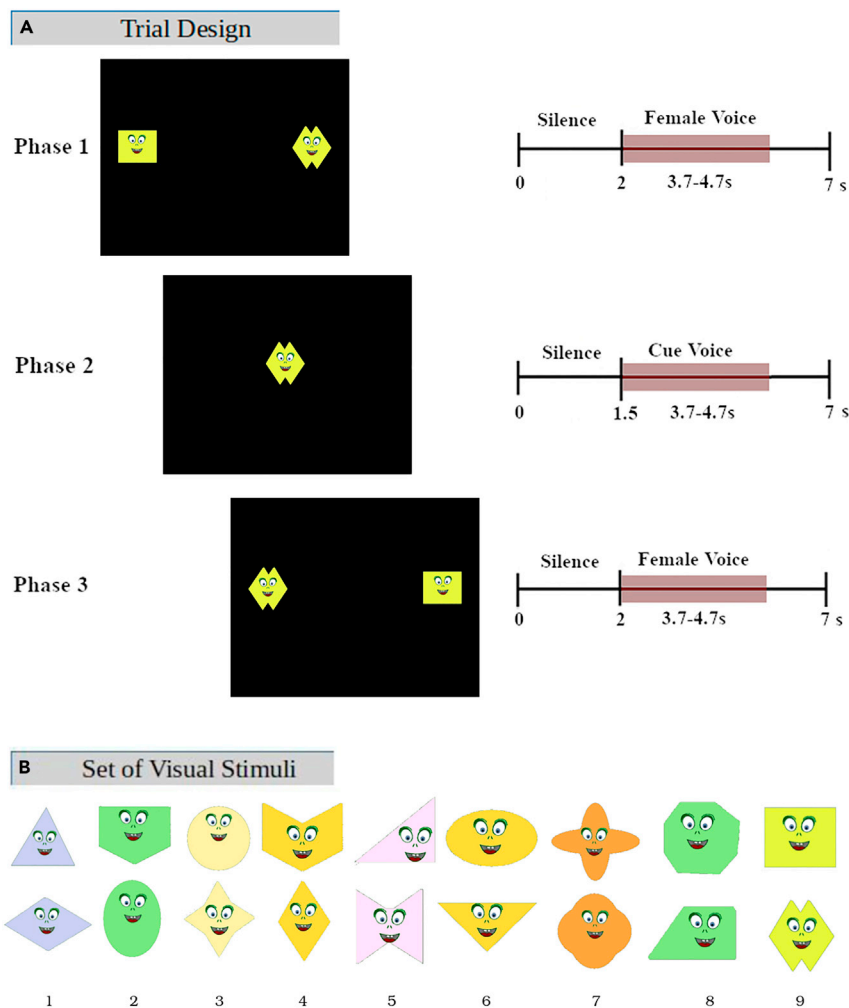


Figure 1. Design of the task presented to adults and infants

(A) Timeline of the three phases of a single trial. In phases 1 and 3 two static faces were presented on the screen. After 2 s of silence a distinct female voice was presented while the faces were still on the screen. In phase 2 one of the two cartoon faces was presented at the center of the screen, accompanied by either a male voice (FMF), the same female voice (FFF), or silence (FSF). The red horizontal bar marks the average length of time the voices were presented.

(B) Set of faces presented across trials. At each trial, a new set of faces and voices was presented.

rectangle face or the diamond face, but not both. In phase 2, one of the faces was presented with a male voice (*cued-face*). In the example of Figure 1, the diamond face has a male voice, so it implies that it is the female voice that was associated with the rectangle face (not A, thus B). Finally, phase 3, which was identical to phase 1, allowed to check whether infants indeed solved the problem: i.e., they should orient toward the *non cued-face* (i.e., the rectangle). We controlled for different explanations of infants' looking behavior in phase 3 by manipulating the evidence provided in phase 2 (Figure 1A). We tested 10-month-olds because at this age infants readily form face-voice associations based on gender (Werker and McLeod, 1989). We also tested a group of adults to control for the dynamics of the inference process.

We measured the orientation time ratio to each of the two faces in phase 3 and compared this ratio across the conditions defined by the evidence provided in phase 2. We also considered the pupil diameter during critical phase 2 when infants must resolve the ambiguity. Pupil diameter is sensitive to cognitive processes (Eckstein, et al., 2017), arousal, and cognitive effort inducing pupil dilatation (Mathôt, et al., 2018). This measure has been proposed as a more sensitive measure for tracking infants' cognitive engagement than looking time (Jackson and Sirois, 2009). We propose several exploratory analyses using this index to look for higher cognitive load in the disjunctive inference condition and eventually reveal whether infants

were actually considering two hypotheses rather than focusing on only one possibility as discussed by Leahy and Carey (2020).

RESULTS

We tested three different conditions in three different groups in infants but within-subject in adults. The first condition was described above and corresponded to the disjunctive inference condition or *Female-Male-Female* (FMF) condition. The disjunctive inference framework predicts that, in this condition, infants would look less at the cued-face in the third phase of the trials when they hear the female voice and orient to the other face, since the cued-face was associated with a male voice in phase 2. However, in this design, the cued-face was presented longer relative to the non-cued-face, a difference equivalent to the duration of the second phase. Thus, in phase 3, the non-cued-face might have appeared more novel and attracted the infants' attention. To control for this potential bias, no voice was presented during phase 2 in a second condition (Female-Silence-Female [FSF] condition). If infants were sensitive to the partial novelty of the non-cued-face, they would look more at the non-cued-face in the third phase as in the FMF condition. In contrast, if partial novelty itself did not drive the infants' looking pattern, no preference for any face should be observed. Finally in a third condition (*Female-Female-Female* [FFF] condition), we tested a simple associative task in which the same female voice was presented in phases 1, 2, and 3. This gave participants direct evidence that the female voice corresponded to the cued-face. Thus, we expected that the subjects would look longer at the cued-face in phase 3, as this face was previously associated with the same female voice in phases 2 and 3. Finally, a group of infants participated in a last condition in which we eliminated phase 1 (*Male-Female* [_{MF}] condition). Indeed, in the FMF condition infants can ignore phase 1 (Jasbi et al.'s explanation of Cesana-Arlotti et al.'s results), wait for phase 2 during which they gather evidence on the voice-face match, then in phase 3, use this knowledge to attribute the new voice to the other (non-cued-) face. In this case, both the FMF and _{MF} conditions should result in a similar gaze pattern in infants. By contrast, if infants consider and rely on the two alternatives raised by phase 1 (i.e., that the voice belongs to the diamond or to the rectangle), we would expect to observe a chance performance in the _{MF} condition, since omission of the first phase would hinder the disjunctive inference process. In all the groups and conditions, the sets of faces and voices were varying across trials.

Adults rely on cue voice in phase 2 to disambiguate the female face

In adults, the average ratio of gaze at the cued-face was 0.35 ($t(13) = -2.30$, $p = 0.039$) for the FMF condition; 0.56 ($t(13) < 1$) for the FSF condition, and finally 0.72 ($t(14) = 2.49$, $p = 0.026$) for the FFF condition (Figure 2). A one-way within-subject ANOVA revealed a significant interaction across conditions ($F(2,42) = 6.38$, $p = 0.0038$). Planned paired comparisons revealed a significant difference between FMF and FFF conditions ($t(27) = -3.32$, $p = 0.0026$, Cohen's $d = 1.23$) but not between FFF and FSF conditions ($t(27) = -1.45$, $p = 0.16$). Contrasting the FMF and FSF conditions resulted in a Cohen's effect size of $d = 0.94$ and a significant difference ($t(27) = 2.40$, $p = 0.024$). Using cluster-based permutation test, we observed a significant cluster in the comparison FMF versus FFF (2.43–6.7 s from the face onset, $p = 0.0087$), as well as in FMF versus FSF comparison (2.65–5.3 s from the face onset, $p = 0.025$). Adults thus took into account the voice presented during phase 2 to correctly match voice and face in phase 3, whereas they remained at chance when there was no voice in phase 2.

The gaze pattern of 10-month-old infants is consistent with disjunctive inference

Unlike the adults, each infant participated in a single condition. The average gaze ratio at the cued-face was 0.35 (SD = 0.11; $t(14) = 5.39$, $p < 0.0001$, two tailed t test) for the FMF condition; 0.52 (SD = 0.11, $t(17) < 1$) for the FSF condition, and 0.54 (SD = 0.082, $t(12) = 2.45$, $p = 0.030$) for the FFF condition. A one-way between-subjects ANOVA comparing the performance of the three groups was significant ($F(2, 43) = 17.26$, $p < 0.001$). Planned paired comparisons revealed a significant difference between FMF and FFF ($t(26) = -5.62$, $p < 0.0001$, Cohen's effect size $d = 2.15$), FMF and FSF conditions ($t(31) = 4.44$, $p < 0.001$, Cohen's $d = 1.55$), and a trend for FFF versus FSF ($t(29) = -2.014$, $p = 0.053$, Cohen's $d = 0.75$). Cluster-based permutations identified two significant clusters in the comparison FMF versus FFF (3–4.3 s, $p = 0.020$ and 5.01–5.87 s from the face onset, $p = 0.022$) and one cluster in the contrast FMF versus FSF (2.8–3.9 s from the face onset, $p = 0.028$) (Figures 3A and 3B).

Stating the problem in phase 1 is used by infants to form a disjunctive inference

In the last group (_{MF} condition), the average gaze ratio and the gaze dynamics remained at the chance level: 0.51 (SD = 0.073, t test $t(22) < 1$), a pattern significantly different from the FMF group ($t(30) = 5.4$,

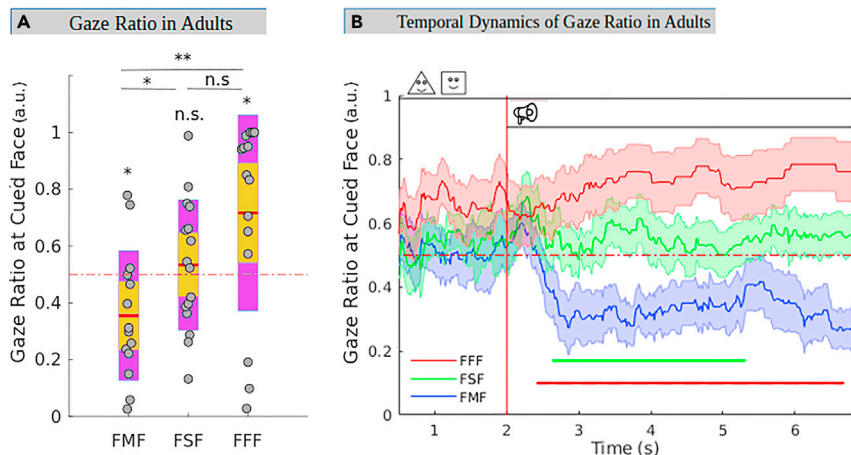


Figure 2. Observed gaze ratio at the correct face in the adult group across conditions

(A) Overall gaze ratio at the *cued-face* during the voice presentation in phase 3 in the adult group. Each circle represents a within-subject average in each condition (*FMF*, *FSF*, and *FFF*) * $p < 0.05$, ** $p < 0.005$. Yellow areas of each bar represent the 95% confidence interval, and the violet areas represent the extent of 1 standard deviation.

(B) Temporal dynamics of gaze ratio at the *cued-face* in each within-subject condition during phase 3. Zero corresponds to the onset of the faces, and the red vertical line indicates the voice onset. Shaded areas represent the standard errors. Horizontal colored lines represent the time span at which the ratio of gaze at *cued-face* in *FMF* was significantly diverging from *FFF* (red) and *FSF* (green), respectively, based on cluster-based permutation tests.

$p < 0.0001$, Cohen's $d = 1.88$, Figures 3C and 3D). In short, the comparison of the gaze dynamics across groups suggests a significantly diverging pattern for the *FMF* condition in contrast to all other conditions.

No evidence for the single belief scenario

One way to test the single belief hypothesis is to consider that, if infants attributed the female voice to one of the two faces in phase 1, they probably looked longer to this face (higher gaze ratio) and thus should be surprised (larger pupil dilation) if this preferred face was subsequently associated with the male voice in phase 2. Hence, we performed a correlation analysis between the gaze ratio to the (to-be-) *cued-face* in phase 1 and pupil dilation in phase 2. We found no significant correlation in any of the groups (Pearson correlation, *FMF*: $r(52) = 0.15$, $p = 0.27$, *FFF*: $r(51) = -0.20$, $p = 0.16$, *FSF*: $r(53) = 0.023$, $p = 0.90$, Figure 4A). Although the interpretation of a null effect is limited, we did not observe any evidence in favor of the formation of a single belief in phase 1 in any of the conditions. Similar results were observed in the adult group.

Second, we examined the relation between pupil dilatation in phase 2 and gaze pattern in phase 3. The hypothesis we followed was that in the *FMF* condition, the process of elimination induced a cognitive load that should be traceable in the pupil dilation, i.e., the more engaged a subject is in inferring the correct pairing through the elimination of the face-voice pair of phase 2, the more they should orient to the other face in phase 3. This was indeed the case in the *FMF* condition (Pearson correlation, *FMF*: $r(58) = -0.28$, $p = 0.030$) but in none of the other conditions (Pearson correlation, *FFF*: $r(59) = -0.013$, $p = 0.92$; *FSF*: $r(86) = 0.063$, $p = 0.56$, *_MF*: $r(109) = 0.033$, $p = 0.73$; Figure 4B). The same analyses in the adult group only resulted in a weak trend at the *FMF* condition (Pearson correlation, *FMF*: $r(55) = -0.23$, $p = 0.085$, *FSF*: $r(60) = 0.09$, $p = 0.49$, *FFF*: $r(59) = -0.10$, $p = 0.44$). The lack of a consistent effect in the *FMF* condition in this group might be due to the relative ease of the negation/disambiguation process in adults, in contrast to the 10-month-old infants.

DISCUSSION

We tested 10-month-old infants in a disambiguation task based on face-voice associations. Our goal was to determine whether preverbal infants were capable of making a disjunctive inference. Infants' behavior was significantly different in the *FMF* condition compared with all control conditions, suggesting that they correctly deduced that the female voice should be paired with the other face given the negative evidence in phase 2. We first demonstrated that infant gaze pattern is closely in line with the pattern observed in

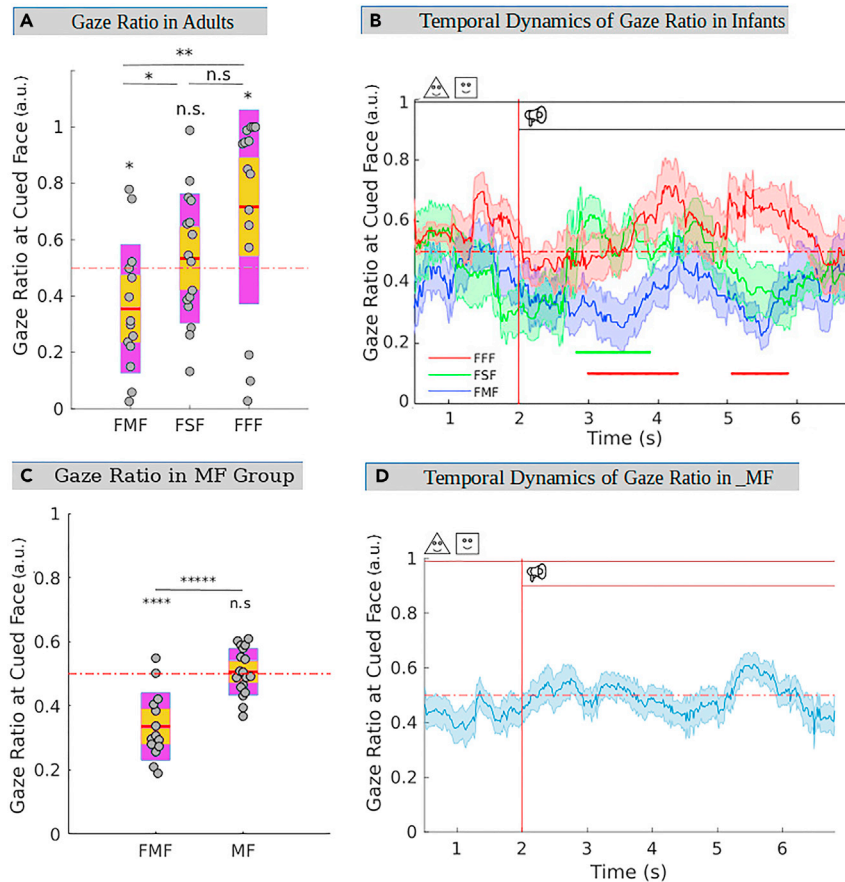


Figure 3. Observed gaze ratio of infants at the correct face across conditions

(A) Overall gaze ratio at the *cued-face* at the *cued-face* in 10-month-olds in the three *FMF*, *FSF*, and *FFF* groups during voice presentation. $**p < 0.005$. Yellow areas of each bar represent the 95% confidence interval, and the violet areas represent the extent of 1 standard deviation.

(B) Temporal dynamics of gaze ratio at *cued-face* at phase 3 in the three groups. Chance level at 0.5, time is measured from face onset. Female voice onset is at 2 s after face onset. The shaded areas mark the standard errors. Horizontal colored lines represent the time span(s) at which the ratio of gaze at *cued-face* in *FMF* was significantly diverging from *FFF* (red) and *FSF* (green), respectively, based on cluster-based permutation tests.

(C) Overall gaze ratio at the *cued-face* at *FMF* condition (same data as in the plot A) in contrast to the *_MF* condition. $****p < 0.000005$.

(D) Temporal dynamics of gaze ratio at *cued-face* in *_MF* group. Chance level at 0.5, the vertical red line marks the voice onset. The shaded area marks the standard error.

adults on an identical task. We then ruled out several alternative explanations: (1) that their performance was related to low-level salience effects (*FSF* control); (2) that they discarded the *cued-face* because it was more familiar (*FFF* control), and crucially (3) that they waited until phase 2 to decide, and thus ignored phase 1 ambiguity (*_MF* control). Furthermore, pupil diameter analyses suggested that there was a specific cognitive load in phase 2 of the *FMF* condition, which was predictive of successful inference in phase 3.

The behavior of the infants was very similar to that of adults, who are capable of disjunctive inference. Although the overt task was lacking in complexity for adults, it is noteworthy that the pupil dilation component that marks the implicit calculation shares a similar temporal dynamics in infants and adults, supporting the hypothesis of a common mechanism at both ages.

Our main argument of disjunctive inference reasoning in infants crucially relies on the difference between the *FMF* and *_MF* conditions, with a rather surprising result. In the *_MF* condition, the male voice presented in phase 2 was *a-priori* sufficient to understand that the female voice in phase 3 should be paired with the

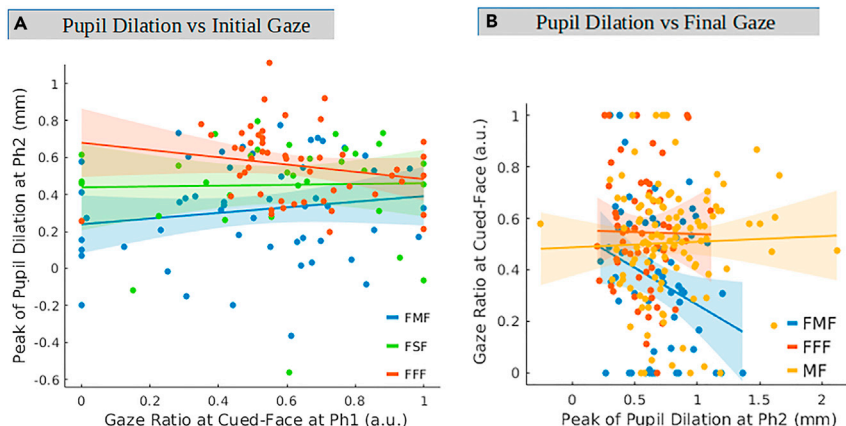


Figure 4. Interaction between the pupil dilation at the cue phase and performance in the infant groups

(A) Correlation between the gaze ratio at the (to-be-) *Cued Face* in phase 1 and the peak of pupil dilation in phase 2. None of the correlations was significant, providing no evidence for a surprise-evoked pupil dilation due to the attribution of the female voice to one of the faces in phase 1. The shaded areas mark the 95% confidence intervals.

(B) Correlation between the pupil dilation peak in phase 2 and the gaze ratio at the *cued-face* in phase 3 in the three conditions FMF, FFF, and *_MF* in infants. The shaded areas mark the 95% confidence intervals. The correlation was only significant for the FMF condition: a larger pupil dilation in phase 2 led to a higher gaze ratio to the correct face (*non-cued face*).

novel face; However, infants randomly oriented to either face in this condition; a behavior significantly different from that of the *FMF* condition. The *_MF* group failure reveals the central role of uncertainty in phase 1 in drawing infants' attention to the relevant problem at hand. This uncertainty might have promoted an attentional engagement and a better encoding of the cartoon face in phase 2, and thus a better individuation of the correct face in phase 3. This explanation is compatible with the correlation between the pupil dilatation in phase 2 and the performance at phase 3, signaling cognitive load and attentional engagement and subsequent correct orientation (Figure 4). Thus, forward reasoning appears easier than retrieving information backward in the working memory when the problem emerges.

As discussed earlier, it has been proposed that, in a situation of ambiguity such as phase 1, infants might form only a single arbitrary belief about the speaking face. This belief is subsequently confirmed or rejected in phase 2. However, we did not observe any interaction between the pupil diameter at phase 2 and an arbitrary preference for one of the faces at phase 1 to support this interpretation. Although negative evidence is a weak argument, our measure was sensitive enough to show a correlation between pupil dilatation in phase 2 and correct orientation in phase 3. This correlation was only evident in the *FMF* condition, revealing a higher cognitive load in situations where disambiguation between two alternatives is required.

Pupil dilation is not specific to a precise cognitive operation, and it can also be induced by surprise should infants assume a face-voice pairing that was subsequently refuted in the second phase. This outcome was postulated in the single hypothesis model (Leahy and Carey, 2020). In this model infants should be correct approximately half of the trials; however, we did not observe any such effect on the pupil dilation. We therefore regard disjunctive inference as the most parsimonious explanation that accounts for these results and those of the *MF* condition. This conclusion, in agreement with Cesana-Arlotti et al. (2018), for the first time to our knowledge, addresses the crucial role of ambiguity during the inference process.

These results contribute to the accumulating evidence showing that preverbal human infants have access to a set of abstract symbols and logical operators to help them to understand their environment (Kabdebon and Dehaene-Lambertz, 2019; Marcus et al., 1999; Cesana-Arlotti, et al., 2020). It raises the question of whether logical operators are abstract amodal concepts or domain specific. Moreover, it questions the previously assumed role the linguistic system plays in logical operations. Do these two systems bear on the same neuronal resources or are symbolic and logical operations general properties of the human brain that are expressed in different domains (Monti, et al., 2012; Monti and Osherson, 2012)? In adults, reasoning on mathematical objects or on factual knowledge recruits different regions, covering a dorsal frontal-parietal pathway, contrasting with the classical perisylvian linguistic network (Monti, et al., 2012; Monti and Osherson, 2012; Amalric and Dehaene, 2019;

Reverberij, et al., 2007). The adult sophisticated mathematical abilities are grounded in the parietal areas (Amalric and Dehaene, 2016), building upon a set of proto-mathematical abilities observed since birth and associated with approximate numerosity (Izard, et al., 2009) and probably also simple geometry (Sablé-Meyer, et al., 2021). The type of calculations these regions support rapidly outperform what has been witnessed in non-human animals. This may be due to the facilitated integration of abstract compressed representations (i.e., symbols) with logical operations since early infancy in a dorsal parieto-frontal pathway, which might be distinct from the more ventral temporo-frontal linguistic system. Further brain imaging studies in infants might clarify the commonalities between logical tasks operated in different cognitive domains (e.g., Cesana-Arlotti's paradigm using objects and the design discussed here using social knowledge) as well as the relations with respect to the linguistic system.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Data and code availability
- MATERIALS, DESIGN AND PROCEDURE
 - Experimental model and subject details
- METHOD DETAILS
 - Stimuli
 - Procedure
- QUANTIFICATION AND STATISTICAL ANALYSIS

ACKNOWLEDGMENTS

We specially thank Francesca Gandolfo for her precious efforts in recruiting the infant subjects; Nick Moalem for reading the draft; Luca Bonatti, Marina Nespor, Bahia Guellai, Amanda Saksia, Hanna Marno, and Alan Langus for the fruitful discussions; and all the adults, parents, and infants who participated in the study. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/European Research Council (ERC) grant agreement No. 269502 (PASCAL) (to J.M.) as well as European Union's Horizon 2020 research and innovation program/ERC grant agreement No. 695710 (to G.D.-L.).

AUTHOR CONTRIBUTIONS

M.E. and J.M. developed the study concept. M.E. designed the experiments and collected and analyzed the data. M.E. and G.D.-L. wrote the manuscript. J.M. secured the funding to conduct the study.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 23, 2021

Revised: August 30, 2021

Accepted: September 28, 2021

Published: October 22, 2021

REFERENCES

- Amalric, M., and Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians. *Proc. Natl. Acad. Sci. U S A* 113, 4909–4917. <https://doi.org/10.1073/pnas.1603205113>.
- Amalric, M., and Dehaene, S. (2019). A distinct cortical network for mathematical knowledge in the human brain. *Neuroimage* 189, 19–31. <https://doi.org/10.1016/j.neuroimage.2019.01.001>.
- Bahrack, L.E., Hernandez-Reif, M., and Flom, R. (2005). The development of infant learning about specific face-voice relations. *Dev. Psychol.* 41, 541–552. <https://doi.org/10.1037/0012-1649.41.3.541>.
- Beatty, J., and Lucero-Wagoner, B. (2000). The pupillary system. In *Handbook of Psychophysiology*, J.T. Cacioppo, L.G. Tassinary, and G.G. Berntson, eds. (Cambridge University Press), pp. 142–162.
- Bonatti, L., Frot, E., Zangl, R., and Mehler, J. (2002). The human first hypothesis: identification of conspecifics and individuation of objects in the young infant. *Cognit. Psychol.* 44, 388–426. <https://doi.org/10.1006/cogp.2002.0779>.
- Brookes, H., Slater, A., Quinn, P.C., Lewkowicz, D.J., Hayes, R., and Brown, E. (2001). Three-month-old infants learn arbitrary auditory-visual pairings between voices and faces. *Inf. Child*

- Develop. 10, 75–82. <https://doi.org/10.1002/icd.249>.
- Cesana-Arlotti, N., Martín, A., Téglás, E., Vorobyova, L., Cetnarski, R., and Bonatti, L.L. (2018). Precursors of logical reasoning in preverbal human infants. *Science* 359, 1263–1266. <https://doi.org/10.1126/science.aao3539>.
- Cesana-Arlotti, N., Kovács, Á.M., and Téglás, E. (2020). Infants recruit logic to learn about the social world. *Nat. Commun.* 11, 1–9. <https://doi.org/10.1038/s41467-020-19734-5>.
- Decasper, A.J., and Prescott, P.A. (1984). Human newborns' perception of male voices: preference, discrimination, and reinforcing value. *Dev. Psychobiol.* 17, 481–491. <https://doi.org/10.1002/dev.420170506>.
- Eckstein, M.K., Guerra-Carrillo, B., Miller Singley, A.T., and Bunge, S.A. (2017). Beyond eye gaze: what else can eyetracking reveal about cognition and cognitive development? *Dev. Cogn. Neurosci.* 25, 69–91. <https://doi.org/10.1016/j.dcn.2016.11.001>.
- Fulkerson, A.L., and Waxman, S.R. (2007). Words (but not Tones) facilitate object categorization: evidence from 6- and 12-month-olds. *Cognition* 105, 218–228. <https://doi.org/10.1016/j.cognition.2006.09.005>.
- Ghirlanda, S., Lind, J., and Enquist, M. (2017). Memory for stimulus sequences: a divide between humans and other animals? *R. Soc. Open Sci.* 4, 161011. <https://doi.org/10.1098/rsos.161011>.
- Goren, C.C., Sarty, M., and Wu, P.Y.K. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics* 56, 544–549.
- Hamlin, J.K., Wynn, K., and Bloom, P. (2007). Social evaluation by preverbal infants. *Nature* 450, 557–559. <https://doi.org/10.1038/nature06288>.
- Inhelder, B., and Piaget, J. (1958). An essay on the construction of formal operational structures. In *The Growth of Logical Thinking: From Childhood to Adolescence* (pp. 67–79), A. Parsons and S. Milgram, eds. (Basic Books).
- Izard, V., Sann, C., Spelke, E.S., and Streri, A. (2009). Newborn infants perceive abstract numbers. *Proc. Natl. Acad. Sci. U S A* 106, 10382–10385. <https://doi.org/10.1073/pnas.0812142106>.
- Jackson, I., and Sirois, S. (2009). Infant cognition: going full factorial with pupil dilation. *Dev. Sci.* 12, 670–679. <https://doi.org/10.1111/j.1467-7687.2008.00805.x>.
- Jasbi, M., Bohn, M., Long, B., Fourtassi, A., and Frank, M.C. (2019). Comment on Cesana-Arlotti et al. (2018). *ResearchGate*. <https://doi.org/10.31234/osf.io/g2h7m>.
- Johnson, M.H., Dziurawiec, S., Ellis, H., and Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition* 40, 1–19. [https://doi.org/10.1016/0010-0277\(91\)90045-6](https://doi.org/10.1016/0010-0277(91)90045-6).
- Jordan, K.E., and Brannon, E.M. (2006). The multisensory representation of number in infancy. *Proc. Natl. Acad. Sci. U S A* 103, 3486–3489. <https://doi.org/10.1073/pnas.0508107103>.
- Kabdebon, C., and Dehaene-Lambertz, G. (2019). Symbolic labeling in 5-month-old human infants. *Proc. Natl. Acad. Sci. U S A* 116, 5805–5810. <https://doi.org/10.1073/pnas.1809144116>.
- Leahy, B.P., and Carey, S.E. (2020). The acquisition of modal concepts. *Trends Cogn. Sci.* 24, 65–78. <https://doi.org/10.1016/j.tics.2019.11.004>.
- Marcus, G.F., Vijayan, S., Bandi Rao, S., and Vishton, P.M. (1999). Rule learning by seven-month-old infants. *Science* 283, 77–80. <https://doi.org/10.1126/science.283.5398.77>.
- Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>.
- Mathôt, S., Fabius, J., Van Heusden, E., and Van der Stigchel, S. (2018). Safe and sensible preprocessing and baseline correction of pupil-size data. *Behav. Res. Methods* 50, 94–106. <https://doi.org/10.3758/s13428-017-1007-2>.
- Medam, T., Marzouki, Y., Montant, M., and Fagot, J. (2016). Categorization does not promote symmetry in Guinea baboons (*Papio papio*). *Anim. Cogn.* 19, 987–998.
- Mody, S., and 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>.
- Monti, M.M., and Osherson, D.N. (2012). Logic, language and the brain. *Brain Res.* 1428, 33–42. <https://doi.org/10.1016/j.brainres.2011.05.061>.
- Monti, M.M., Parsons, L.M., and Osherson, D.N. (2012). Thought beyond language: neural dissociation of algebra and natural language. *Psychol. Sci.* 23, 914–922. <https://doi.org/10.1177/0956797612437427>.
- Orena, A.J., and Werker, J.F. (2021). Infants' mapping of new faces to new voices. *Child Dev.* 92, e1048–e1060.
- Reverberi, C., Cherubini, P., Rapisarda, A., Rigamonti, E., Caltagirone, C., Frackowiak, R.S.J., and Paulesu, E. (2007). Neural basis of generation of conclusions in elementary deduction. *Neuroimage* 38, 752–762. <https://doi.org/10.1016/j.neuroimage.2007.07.060>.
- Sablé-Meyer, M., Fagot, J., Caparos, S., van Kerkoerle, T., Amalric, M., and Dehaene, S. (2021). Sensitivity to geometric shape regularity in humans and baboons: a putative signature of human singularity. *Proc. Natl. Acad. Sci. U S A* 118, e2023123118. <https://doi.org/10.1073/pnas.2023123118>.
- Standley, J.M., and Madsen, C.K. (1990). Comparison of infant preferences and responses to auditory stimuli: music, mother, and other female voice. *J. Music Ther.* 27, 54–97. <https://doi.org/10.1093/jmt/27.2.54>.
- Werker, J.F., and McLeod, P.J. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Can. J. Psychol.* 43, 230–246. <https://doi.org/10.1037/h0084224>.
- Xu, F., Carey, S., and Quint, N. (2004). The emergence of kind-based object individuation in infancy. *Cognit. Psychol.* 49, 155–190. <https://doi.org/10.1016/j.cogpsych.2004.01.001>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Stimuli presentation	Psyscope Build 70	http://psy.cns.sissa.it/
Occulomotor and Pupil Recording	Tobii T60, Tobii Studio	https://www.tobii.com/
Occulomotor and Pupil Analyses Algorithms	This paper	https://osf.io/cxrm3/
Stimuli	This Paper	https://osf.io/cxrm3/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for data and the scripts should be directed to and will be fulfilled by the lead contact, Milad Ekramnia (Milad.Ekramnia@cea.fr)

Data and code availability

- **Data.** All raw data is available at osf.io/cxrm3
- **Code.** The main scripts are provided at osf.io/cxrm3
- **Other.** Sample stimuli are accessible at osf.io/cxrm3

MATERIALS, DESIGN AND PROCEDURE

Experimental model and subject details

Adults: A group of 16 Italian native speakers (20-27 years, 10 females) were tested on the three conditions (FMF, FSF and FFF) with an identical temporal design. A subject was included in the analyses if s/he passed at least 3 out of 5 of valid trials in each condition. Of 16 subjects, 14 were included in the FMF condition, 14 in the FSF and 15 in the FFF condition. No subject was eliminated because of a looking bias to one side of the screen.

Infants: Four groups of monolingual Italian infants were tested on the FMF, FFF, FSF and *_MF* conditions. Thirty-four infants were tested (14 girls, mean = 10.24 months(m), SD = 0.44) in the FMF group, 29 infants (12 girls, mean = 10.34 m., SD = 0.41) in the FFF group, 23 (13 girls, mean= 9.8 m., SD=0.42) in the FSF and 32 (16 females, mean =10.59 m., SD=0.6) in the *_MF* group, of which 18, 14, 18 and 18 subjects were included for further analyses in the respective conditions. Participants were recruited from Trieste, Italy, through sending invitation letters to a random selection of parents, whose babies fit the age range of the study. The parents signed a consent form prior to the experiment and the ethical committee of the Scuola Internazionale Superiore di Studi Avanzati (SISSA) approved the study.

METHOD DETAILS

Stimuli

Cartoon faces: Many studies have shown that infants prefer faces to other visual stimuli (Goren, et al., 1975; Johnson, et al., 1991), and that social cues, such as eye contacts, improve the task engagement and the subsequent task performance (Hamlin, et al., 2007; Bonatti, et al., 2002). Proposed explanations for these findings include enhanced infant attentiveness to these cues, and/or a higher degree of competence in the social domains. We thus used cartoon faces in response to which infants might be more likely to attribute a voice. All faces were created in GNU Image Manipulator (GIMP) software. v.2.8.1 and shared identical facial features, with no intrinsically gendered feature. The mouth was half opened to suggest that the face might be speaking. No face was repeated for a subject. The colours of the faces were identical within a trial (Figure 1B) but changed from one trial to the next. Each face occupied a region of 320 x 280 pixels on a black screen with a resolution of 1280 by 1024 pixels. When two faces were presented, they were placed symmetrically at approximately 75 pixels from the left and right edges of the screen. When one face was presented,

it was positioned centrally. The faces at phase 3 of the trials were counterbalanced to avoid potential contamination due to a persistent side gaze following.

Voices: Italian-native female and male speakers were recorded uttering the sentences: “Lungo il fiume su andiamo, tutti insieme saltelliamo” (“Let’s go along the river, all jumping together”), in an infant-directed fashion (Werker and McLeod, 1989). The amplitude of the recorded voices was normalized to 65 db using Praat software (v. 5.0.42) and they were presented in stereo on both sides of the screen. The produced sentences varied in duration (3.7 to 4.7 seconds).

Procedure

General design. Participants were tested in a dark room with an eye-tracker mounted on the screen to measure participant’s eye gaze and pupil dilation profile. The experiment started approximately 2 minutes after the subject was positioned in front of the screen, during which the subject passed an eye-position calibration session, implemented by Tobii studio, and a short animation of hallow patterns on the center of the screen to have the attention engaged on the screen and allow pupil adaptation to the ambient light condition.

We tested 4 conditions: the disjunctive inference condition (FMF) along with the 3 other conditions (FSF, FFF, and _MF) which controlled for different aspects of the FMF design. Each infant participated in one condition only (between-subjects design), contrary to adults who attended all conditions except _MF. Each trial (duration = 23 s) was divided into three phases. In all conditions and phases, a short period of silence was implemented before the voice onset to let the pupil diameter adjust to the luminosity of the new display (2s, 1.5s and 2s for phases 1, 2 and 3 respectively). Voice and cartoon faces were never repeated across trials.

FMF Condition: At phase 1 of each trial, two static cartoon faces identical in colour and facial features, but differing in shapes, were presented on the screen (Figure 1A). After 2 seconds a female speaker uttered two sentences. In Phase 2 one of the two faces was randomly selected and presented at the center of the screen (*i.e. the cued-face*). This image was accompanied with a male voice producing the same sentences as in phase 1. Finally, both cartoon faces were presented simultaneously in Phase 3, counterbalanced in side position across trials and accompanied with the same female voice as in Phase 1. We refer to this condition as *Female-Male-Female (FMF)*. We chose female voices as referent to facilitate the infant task of disambiguating the two faces in phases 1 and 3, because infants prefer female voices (Werker and McLeod, 1989; Standley and Madsen, 1990; Decasper and Prescott, 1984) and might be more inclined to identify the associated cartoon face.

FSF Condition: This condition (*Female-Silence-Female*) was similar to the previous one except that no voice was presented during phase 2.

FFF Condition: This condition (*Female-Female-Female*) was similar to the FMF condition, except that the same female voice was presented at each phase.

_MF Condition: This last condition comprised only the last two phases 2 and 3. The voice onset in phase 2 was being delayed from 1.5s to 4.5s relative to the face onset to let the pupils adapt to the luminosity of the single face on the screen prior to the voice onset. Twelve trials were presented to each participant.

Infants’ Procedure: The infants sat on a chair fixed on the lap of their parents, approximately 70 cm (eye-tracker’s allowed range [50cm-80cm]) from a 17” screen, paired with a Tobii eyetracker T60 (60 Hz sampling rate). Parents were wearing opaque sunglasses to prevent any interference with the infant’s performance. An extra infrared emitter was placed behind the eye-tracker to provide an enhanced eye detection quality.

Prior to the experiment, and after the calibration, infants watched a short animation for six seconds to orient their attention toward the screen center and allow pupil adaptation to the light condition of the room. At the beginning of each trial, a jiggling bell was centrally presented for 2.25s to reorient the infant’s gaze to the center of the screen. If it was not sufficient, a few seconds of a Pixar Animation ‘For the Birds’ was shown. Infants’ head orientation was checked via an infrared camera throughout the experiment. Each infant received 9 trials in the specific condition of the group.

Adults' Procedure: Participants received written instruction to overtly look at the face corresponding to the voice. With the head supported on a chin-rest. They sat approx. 60 cm (eye-tracker allowed range 50-80cm) from a screen, paired with a Tobii eye tracker T120 (120 Hz sampling rate). Each participant received 15 trials, 5 trials in each of the *FMF*, *FFF* and *FSF* conditions in a randomized order.

For all groups, infants and adults, the stimuli were presented by Psyscope software, build 70 (<http://psy.cns.sissa.it/>), furthermore the gaze and pupillometry data was being recorded by Psyscope in junction with Tobii studio.

Oculo-Motor measures. Prior to the experiment, all subjects passed a 5-point-gaze calibration implemented by Tobii studio with a tolerance of one missing point in the calibration measure. Gaze coordinates and pupil diameter was processed only from the left eye for all the analyses in all groups. A side bias threshold for each subject was assigned if more than 80% of the valid gazes during phase 1 of the trials fell in one side of the screen. One infant was rejected because of this criterion in each of the *FMF*, *FFF* and *FSF* groups.

In infant groups the first trial was considered as a task familiarization trial and discarded (omission of this trial did not have any significant effect on the overall performances).

To assign the loci of the gaze points at each time bin, the screen was divided into three equivalent regions of interest (RoI), each covering 33% of the horizontal axis with a 20% margin on the lower and upper extents of the vertical axis. In phases 1 and 3, the left and right RoIs were considered valid and in phase 2, only the central RoI. For a trial to be included, participant gaze should be in a valid RoI during 50% (phase 1 and 3) and 80% (phase 2) of the sentence duration., except for the *FSF* condition in which there was no voice during phase 2 and the threshold was thus lowered to 50% in this case. The same thresholds were used for adults and infants. To be included in this study, adult and infant participants should have at least 3 valid trials in the considered condition.

QUANTIFICATION AND STATISTICAL ANALYSIS

Gaze-ratio Analyses: Our main comparison relies on the infant gaze behaviour in Phase 3 and on whether their orientation to one face was affected by the different preceding phases in function of the conditions. Since Phase 3 was identical in all conditions (two faces with a female voice), we defined our measure as the ratio of valid gaze falling within the RoI of the *cued-face* (i.e. the face presented in phase 2) relative to the other face. We expected 1) a lower gaze ratio in the *FMF* and *_MF* conditions (the *cued-face* is rejected since it was paired in Phase 2 with the male voice); 2) inversely a higher gaze ratio in *FFF* condition (the *cued-face* was paired with the same female voice in phase 2); 3) a ratio at chance for the *FSF* condition since no explicit cue was given in phase 2, and participants could assign the voice to whatever face they prefer.

First, we considered the averaged gaze-ratio during the whole sentence presentation in phase 3. A two tailed t-test against chance-level (50%) was performed for each condition in adults and infants. Conditions were further compared in a one-way ANOVA with a three-level condition factor, within-subject in adults and between subjects in infants, followed by a comparison of the conditions two by two. Second, we considered the temporal evolution of the gaze ratio in the conditions of interest. We performed cluster-based permutation tests (Maris and Oostenveld, 2007) comparing the two conditions of interest across all time bins including the silence period. Clusters were constituted by summing contiguous values of unpaired t-values superior to a threshold ($p_{val} < 0.1$). Then the statistic value of the largest cluster obtained in the original data was determined by comparing to the null distribution determined through 4000 permutations.

Pupil dilatation analyses: Pupil diameter is not only affected by the ambient light but also by different cognitive processes through the involvement of the *locus coeruleus* which modulates the pupil dilation. Thus, it has been shown that the magnitude of pupil diameter variation is a relevant index of underlying cognitive operations (Beatty and Lucero-Wagoner, 2000). We propose several exploratory analyses to investigate whether the cognitive load is different in the *FMF* condition relative to others.

The pupil data, obtained using Tobii software, was first subjected to an artifact rejection process, intended to eliminate jumps in the Tobii eye-tracker measures. The data was then baselined on the minimum of the

pupil diameter during the silent phase, prior to the onset of speech, when the eyes adjust to the brightness of the face(s) on screen. The data was then epoched for each phase relative to voice onset.

First, we considered the relation between phases 1 and 2: If infants started a disjunctive inference in phase 1, they formed two incompatible possibilities: the female speaker can be the left face on the screen or the right one. In phase 2, one of these possibilities is dismissed. Alternatively, infants could have randomly assigned the female voice to one of the cartoon faces. In phase 2, this belief was either confirmed or refuted, creating a surprise in the latter case. Because surprise induces a pupil dilatation, their preference for one or the other face during phase 1 might be positively correlated with the pupil diameter (i.e. the more they orient to the (to-be) *cued-face* during phase 1 while listening to the female voice, the more they should be surprised for its association with the male voice in phase 2). Thus, pupil dilation factor during phase 2 combined with the gaze orientation during phase 1 might be a marker of single belief vs disjunctive inference. Therefore, we computed the correlation between the gaze ratio to the face which was subsequently being cued in phase 2 and the pupil dilatation to this face in phase 2. Furthermore, if infants considered only one face during phases 1 and 2, they should be correct on approximately half of the trials, leading to a bimodal distribution of pupil dilation, in contrast to the more even distribution expected in the case of disjunction where cognitive effort remains roughly similar across trials.

Second, we considered the relation between Phase 2 and 3: In the case of the *FMF* condition, infants dismissed one of the two possible faces to keep only one possibility (not A then B). Even-though this is an all-or-none process, i.e. either infants perform the negation or they don't, however the level of engagement in this process will be reflected as the distribution of performance at phase 3. In other words, the more infants were engaged in this inference, the more they should orient to the *non-cued face* in Phase 3 (i.e. lower gaze ratio to the *cued-face*). We thus computed the correlation between pupil dilatation in Phase 2 with the *cued-face* gaze ratio in Phase 3.