The Origin of Logical Concepts

Nicoló Cesana Arlotti

TESI DOCTORAL UPF/ 2015

DIRECTOR DE LA TESI

Dr. Luca Bonatti

Departament de Tecnologies de la Informació i les Comunicacions

Universitat Pompeu Fabra Barcelona

Acknowledgement

I would like to thank Luca Bonatti for his scientific guidance, constant support and endless patience during these five rich and intense years. I'm extremely grateful to all the members of the Center for Brain and Cognition at Pompeu Fabra University for their scientific advice and friendship. I'm grateful to Núria Sebastián Gallés, Albert Costa, Gustavo Deco, Salvador Soto-Faraco and Juan Manuel Toro for having provided the rich scientific environment in which I was introduced to cognitive science research. I'm especially grateful to the members of the RICO group, both past and present, for the long and passionate discussions within and outside of our group meetings: Alexandra Ibañez Lillo, Alveno Vitale, Ansgar Endress, Gisela Pi Casaus, Katya Mehta, Kinga Bohus, Liza Vorobyova, Nuria Ferres Forga and Ryszard Cetnarski. Special thanks to Nuria Ferres Forga, Liza Vorobyova, Kinga Bohus and Ryszard Cetnarski with whom I had the privilege to work on common projects. I'm immensely grateful to Susan Carey for having welcomed me into her lab. Some of the ideas presented in this dissertation originate from the exchanges I had with her and her students Arin Tuerk, Roman Feiman and Shilpa Mody. I want to thank Jean-Remy Hochmann and Miguel Burgaleta for their useful comments on preliminary drafts of this thesis. I want to dedicate a special thanks to Sandro Zucchi, who ignited my interest in research and prepared me for its asperity. My deepest gratitude goes to Mariella and Ivan for their unconditional love and trust. Finally, I dedicate this work to Olga. Without her, this dissertation wouldn't have been possible.

The research reported in this dissertation was supported by a grant for universities and research centers for the recruitment of new research personnel (FI-DGR 2011) and by a grant for research stay abroad (BE-DGR 2012) offered by the Catalan Government (AGAUR).

Abstract

From a developmental perspective, propositional logical reasoning has been often characterized either as the peak of cognitive development or as an innate property of human cognition. Our ignorance of the origin of logical thinking in part depends on its primary source of evidence. Overt linguistic logical behavior does not begin to arise before the second birthday. As a result, the first 2 years of human development remain completely uncharted by psychologists of logical reasoning.

The aim of this dissertation is to contribute to the beginning of the exploration of preverbal logical abilities. We developed methodology based on entirely non-linguistic logical tasks, implicit measures of scene processing such as cumulative looking time in a violation of expectancy paradigm, and automatic eye-tracking procedures.

The joint use of these techniques offers preliminary evidence that basic logical representations might be tracked at least from the beginning of the second year of life.

Resumen

Desde una perspectiva del desarrollo, el razonamiento lógico proposicional se ha considerado como el cénit del desarrollo cognitivo o como una propiedad innata de la cognición humana. Nuestra ignorancia con respecto al origen del pensamiento lógico depende, en parte, de su fuente principal de evidencia. En efecto, el comportamiento lógico lingüístico no se hace evidente antes del segundo cumpleaños. Como resultado, los primeros dos años de desarrollo humano se mantienen completamente inexplorados para la psicología del razonamiento lógico. El objetivo de esta tesis es contribuir al inicio de la exploración de las capacidades lógicas preverbales. Para ello hemos desarrollado nuevos métodos basados en tareas lógicas completamente no lingüísticas, incluyendo medidas implícitas de procesamiento de escenas como el tiempo de mirada acumulado en un paradigma de violación de expectativas, así como procedimientos automáticos de captura de movimientos oculares (eye-tracking). El uso combinado de estas técnicas ofrece evidencia preliminar de que el rastro de las representaciones lógicas básicas podría remontarse al menos desde el comienzo del segundo año de vida.

Preface: on the ontogenesis of logical concepts

From Jean Piaget we inherited the view that logical reasoning is a late result of cognitive development. Piaget held that the capacity of reasoning deductively over hypotheses begins to emerge at the age of 11 years with the Formal Operational stage, what he took to be the peak of cognitive development (Piaget, 1960). According to the Swiss psychologist, the propositional operations of (material) implication, disjunction, incompatibility, etc., are constructed when the scope of young adolescents' reasoning is extended from objects, their collections and their relations to propositions (Piaget, 1953). The transition from a reasoning centered on objects to one focused on propositions is, for Piaget, related to a shift of interest from actual facts to hypothetical possibilities. Simplifying Piaget's theory, the operators of propositional logic are the result of a slow process of acquisition of several new kinds of representation that start from sensory-motor ones and culminate in abstract logical concepts, in a time scale of years.

Diametrically opposite is the position of Jerry Fodor. For Fodor, the logical concepts with which we are endowed are not learned at all. This conclusion readily follows from its famous and provocative argument that all primitive concepts (in the specific sense of concepts with no definitorial structure) are not learned (Fodor, 1975); (see Carey (Carey, 2014), Laurence & Margolis (Margolis & Laurence, 2011) for critical discussions of the argument; a new argument, aiming to an a priori refutation of concept learning tout court (of both primitive and complex concepts), has been more recently proposed by Fodor (Fodor, 2008); see Margolis & Laurence (Margolis & Laurence, 2011) and Rey (Rey, 2014) for a discussion of this second argument).

Briefly, Fodor argues that learning a concept by hypothesis confirmation (HC) consists of forming new concepts by formulating and confirming hypothesis regarding its definitorial structure. No primitive concept is learned by HC, since the hypothesis regarding its structure will already contain exactly that concept. But, for Fodor, all learning is HC. So primitive concepts that lack a definitorial structure are not learned.

If primitive concepts are not learned, then it is highly probable that logical concepts are not. It is indeed very probable that logical concepts cannot be expressed in terms of non-logical ones (though some of them can be expressed in terms of other logical concepts, as exemplified by De Morgan's laws). As other primitive concepts, logical concepts are then either functioning at birth, become functioning as the result of an innately determined process, or are acquired through some non-cognitive biological process (Fodor, 1975, 2008). Therefore, and contrary to Piaget, a consequence of Fodor's argument is that logical concepts are not the result of cognitive development at all.

Although the strength of its dialectic has often been acknowledged, Fodor's argument for the innateness of primitive concepts has received several critical responses. Some critics have objected that not all learning is a matter of explicit hypothesis testing (Carey, 2009, 2014; Margolis & Laurence, 2011). Others have pointed to the possibility of learning primitive concepts by testing hypothesis that do not rest on the concept definitorial structure (Block, 1986; Carey, 2009, 2014; Laurence & Margolis, 2002).

The replies cited above question the soundness of Fodor's argument, but none specifically address the question of whether and how logical concepts are learned. Thus, we think that, at the current stage of the debate, the ontogenesis of logical concept is entirely open to empirical investigation. Specific hypotheses regarding the development of logical reasoning have to be advanced and tested.

One source of evidence of logical representations is overt linguistic behavior. For example, some uses of natural language words such as the English expressions “not” and “either…or” are probably evidence of representation of the truth-functional operators of negation and disjunction, respectively (Laurence R. Horn, 1989; Jennings, 1994). As a result, one window into the development of logical representations is the acquisition of logical words.

At the current stage, results from language acquisition suggest that linguistic behavior explainable with logical representations begin to emerge between the age of 2 and 4 years. For example, studies argue that young children (though not younger than 3 years) understand the linguistic expressions corresponding to disjunction in their language as expressing inclusive-disjunction, rather than exclusive-disjunction (Crain & Khlentzos, 2010). Researches on negation-related words production suggest that, although the English word “no” is one of the first linguistic expressions to appear in production, episodes of production with a denial function (e.g. the negation of a proposition expressed by a previous assertion) are supposed to begin in the multi-word speech stage (see Dimroth (Dimroth, 2010) for a survey of the literature). Moreover, two recent studies found that evidence of comprehension of negation compatible with a truth-functional interpretation is found only after the second birthday (2 years and 5 months, (Austin, Theakston, Lieven, & Tomasello, 2014); 3 years and 5 months (Nordmeyer & Frank, 2014).

In summary, investigations on the acquisition of words traditionally associated with logical operators suggest that their logical use and understanding emerge only after the age of two. What, then, about logical concepts acquisition? Are logical representations the result of training in logical words and public language in general? Or else are some logical representations in place before the acquisition of logical words, having perhaps a role in the acquisition of that piece of language? That is, are there *preverbal* logical concepts?[1]

Prima facie there seem to be several possibilities for the relationship between language acquisition and logical reasoning development. Preliminarily we can contrast two very broad alternatives. Logical concepts acquisition might be a product of language learning (see Fodor (Fodor, 1975) and Crain & Khlentzos (Crain & Khlentzos, 2010) for a critical discussion of this possibility). Alternatively, some logical concepts might predate the acquisition of logical words and possibly play a decisive role in it (see Crain & Khlentzos (Crain & Khlentzos, 2010) for a proposal of this type). A direct way to begin to explore those alternatives is by looking for logical representations that are in place before the words with logical meaning are acquired. That is, it is looking for preverbal logical concepts.

Based on the current potential estimation of the early appearance (of logical use and understanding) of logical words, preverbal logical concepts should be searched before the second birthday. Although, to our knowledge, there are no studies that provide substantial evidence of logical reasoning in the second year of life or before, there are cues from the field of word learning that are suggestive of deductive abilities (see sections 1.3.1 and 1.3.2 for a discussion of those studies).

---

[1]We will use the qualification "preverbal" applied to logical concept to specifically mean a logical concept that is present before the attested onset of logical use or understanding of related logical words.

Thus, the very possibility that human infants are endowed with preverbal logical concepts, and in general the development of logical reasoning, demand investigation.

This dissertation aims to offer a preliminary contribution to the investigation of the origin of logical concepts. In a series of experiments we looked for evidence of preverbal logical representations and their independence of language.

Table of contents

List of figures

List of tables

1. GENERAL INTRODUCTION

1.1. Logical concepts and logical inferences

Not much is known of preverbal logical concepts. At the current stage of investigation hypotheses regarding the nature of preverbal logical representations cannot but have a tentative character. A parallel with the literature on preverbal numerical abilities can suggest the extent to which the nature of preverbal logical concepts is unknown. The two systems of preverbal numerical representations that are at the origin of the numerical cognition turned out to be rather different from the concepts studied in formal arithmetic (Carey, 2009; Feigenson, Dehaene, & Spelke, 2004).

Nevertheless, we find it reasonable to assume as a working hypothesis a similarity between preverbal logical concepts and the subject matter of formal logic. Thus, we planned our research to look for representations with (at least) some of the properties of basic logical concepts of formal logic. The extent to which the analogy with formal logic is fitting (i.e. what properties, if any, preverbal representations share with formal logic concepts) is an empirical question that can hardly be exhausted by the present investigation.

Having in mind the standard treatment in formal logic, the preverbal equivalent of a logical sentential connectives (like the sentential operators usually associated with English words "not", "and", "either… or", "if… then", etc.) can be tentatively described as a mental representation with a set of *syntactical rules* of composition with other representations, and having *a truth-function* as semantic value and/or entitling a set of *inferential rules*.

A PLC (preverbal logical concept) might have a proprietary mental symbol and be part of a system of representations (like sentential operators are part of a language). It might have syntactic combinatorial rules through which it is combined with other representations of the system, to form compound representations. For example, the mature concept EITHER… OR can be combined with the representations JOE IS INNOCENT and HE WILL GO TO PRISON to form the compound EITHER JOE IS INNOCENT OR HE WILL GO TO PRISON.

A PLC might be associated with a distinctive truth-function, a function from truth values (e.g. true, false) to truth values. This function might determine whether a compound representation (formed by the PLC with other representations) is true or false: the truth value of the compound is a function of the truth values of its constituents. For example, the truth-function associated with mature inclusive EITHER… OR is such that, for any pair of representations P and Q, the compound representation EITHER P OR Q is true if one of P and Q is true, otherwise it is false. Therefore, EITHER JOE IS INNOCENT OR HE WILL GO TO PRISON is false only in the case that Joe is guilty but he does not go to prison.

A PLC might be connected to a specific set of inferential rules (Block, 1986; C Peacocke, 1992), among many others). Those rules might determine the set of inferential transitions that the concept supports in deduction. Roughly, being equipped with a PLC might causes one to perform the inferential transactions individuated by the rules, when one engages in deduction. Some of these inferential rules might be the result of the interaction of multiple PLC s. For example, mature EITHER… OR and NOT together support an inferential rule called disjunctive syllogism (also known as Modus Tollendo Ponens or disjunction elimination). The rule can be summarized with this formula:

EITHER P OR Q, NOT P ⊢ Q

The formula can be taken to mean that a representation with the form EITHER P OR Q and a representation with the form NOT P together entail Q (P and Q are variables for sentential representations). So, according to the disjunctive syllogism, EITHER JOE IS INNOCENT OR HE WILL GO TO PRISON together with JOE IS NOT INNOCENT entails HE WILL GO TO PRISON.

In summary, we take as a working hypothesis that also PLC might have syntactic combinatorial rules, truth-functional meanings and inferential roles, of the types indicated above. Of those properties, in the current research we focused on the third: the inferential role.

We take evidence of the execution of inferences in accordance with the inferential role of a concept as evidence of the possession of that concept, since the possession of the concept is taken as a causal explanation of the execution of that inferences (Block, 1986; Fodor, 1975, 2004; Peacocke, 1992, 2004). In this dissertation we tackle the question of whether infants and toddlers, who have not yet mastered logical words, have representations with the inferential role of a logical concept. Specifically, we focused on the inferential rule disjunctive syllogism as evidence of preverbal logical disjunction and negation. We used the disjunctive syllogism as a case study, to take a first step in the investigation of the very existence and development of preverbal logical reasoning.

1.2. The disjunctive syllogism

The inferences supported by the disjunctive syllogism consist of three representations: the disjunctive premise, the negative premise and the conclusion.

i) *Either* Joe is innocent *or* he will go to prison (disjunctive premise).

ii) Joe is *not* innocent (negative premise).

iii) Therefore, he will go to prison (conclusion).

Informally speaking, the function of the disjunctive premise is to represent two hypotheses as jointly exhaustive, since the disjunction means that at least one them is true. The role of the negative premise is to falsify one of the hypotheses. The conclusion is a valid consequence of the premises, it cannot be false if the premises are true. The disjunctive syllogism prescribes to draw the valid conclusion (iii) in response to the representations (i) and (ii).

Although the disjunctive syllogism is an inferential rule that an inference satisfies or do not, in order to simplify the exposition we will use "disjunctive syllogism" to refer either to the rule or to an inference that satisfies it; the context will clarify of which of the two the expression refers.

It is important to realize that conformity to the disjunctive syllogism can be alternatively explained by two importantly different sets of logical concepts. One is the pair formed by NOT and inclusive EITHER… OR, the other is the pair formed by NOT and exclusive

EITHER… OR. Both sets of concepts support the disjunctive syllogism. However, the two sets are very different since they have very different expressive powers. The set of NOT and inclusive EITHER… OR is indeed functionally complete. That is, it is sufficient to express all truth-functions and it therefore has the expressive power of the whole propositional logic. The set of NOT and exclusive EITHER… OR is instead not functionally complete.

However, since both sets of connectives support the syllogism, this ambiguity has no relevance for the studies that we will present in this dissertation. In what follows, we will generically speak of OR and of the concept of disjunction without resolving the ambiguity between the inclusive and exclusive disjunction.

One of the reasons for why we decided to study the development of the disjunctive syllogism and not of other rules is that it is a very simple inferential rule (simple in terms of the representational abilities it requires) that explains a fundamental reasoning strategy: reasoning by exclusion. Reasoning by exclusion consists of reaching a conclusion based on evidence that rules out its alternatives (Stalnaker, 1987). This very natural way to reason can be explained in terms of the disjunctive syllogism. Reasoning by exclusion can be formalized as an instance of the disjunctive syllogism. That is, if P and its alternatives are represented disjunctively, the negation of all alternatives will result, by disjunctive syllogism, in the conclusion of P.

Reasoning by exclusion has become an object of investigation in developmental and comparative psychology, where it has been explicitly related with the disjunctive syllogism; in the following sections we present a circumscribed discussion of this literature.

1.3. Clues of early logical reasoning

1.3.1. Disambiguation of novel name referent

The possibility that young children, and perhaps even infants, are capable of elementary logical inferences is under debate in the study of early word leaning (Halberda, 2003, 2006; Mather & Plunkett, 2011). In this and the following sections we will review results and theoretical accounts that ground this debate. Finally, we will try to provide a preliminary evaluation of the actual support offered to the attribution of logical representation to toddlers and infants.

Learning the meanings of an unknown language by observation of speakers' overt linguistic behavior appears to be a difficult enterprise, even for adults who have mastered their own native language (Quine, 1960). For example, to grasp the denotation of an unknown name might require the listener to deal, somehow, with multiple hypotheses, since many possible meanings are compatible with the way a name is publicly used (Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992; Markman & Wachtel, 1988). Yet infants seem able to solve the ambiguity involved in designation by unknown names, since in the first years of life the their vocabulary size rapidly increases (Carey, 1978). Researchers have worked for decades to clarify how infants and children recognize the referent of novel words, for example by proposing biases that constrain their hypotheses regarding the meaning of words, and by studying the way they can benefit from communicative cues offered by competent speakers.

A type of referential ambiguity that developmental psychologists have systematically investigated is that resulting from the utterance of a novel name in the presence of one novel object and one familiar object (or, sometimes, multiple familiar objects). One object is familiar and its name is known. The other object is novel and its name unknown. The linguistic context suggests that the novel name is pointing at one of the objects in the scene (e.g. the name is embedded in a phrase like "look at the _ !"). However, it is insufficient to determine its exact referent: the novel name might refer either to the novel object or to the known object. Numerous studies reported that adults as well as young children exhibit a tendency to take the novel name to refer to novel object, either by manually selecting the novel object or by orienting their attention toward it, in response to the novel name (Bion, Borovsky, & Fernald, 2013; Golinkoff et al., 1992; Halberda, 2006; Horst & Samuelson, 2008; Markman & Wachtel, 1988; Mather & Plunkett, 2012). This behavior has been occasionally documented even in infants and toddlers in their second year of life (Halberda, 2003; Houston-Price, Caloghiris, & Raviglione, 2010; Markman, Wasow, & Hansen, 2003; Fei Xu, Cote, & Baker, 2005)

Thus, experimental evidence suggests that adults, young children and possibly infants have a tendency to take novel names to refer to novel name-unknown objects rather than familiar name-known objects. We will refer to this tendency as *novel name disambiguation*. Although novel name disambiguation is a robust phenomenon, it is not clear what the processes underlying it are. Several competing explanations for that response have been proposed, and are currently the object of lively research.

According to Markman and Watchel (Markman & Wachtel, 1988) *Mutual Exclusivity* is the word learning bias that (category) names denote mutually exclusive categories. The bias implies that distinct category names do not refer to the same object. Although Mutual Exclusivity can be overridden by clear evidence to the contrary (as in the case of superordinate or subordinate categories, and synonyms), it results in a default tendency to exclude that name-known objects may be the referent of novel names, in favor of name-unknown objects.

The *Pragmatic Accounts* ground novel name disambiguation in social communicative competence, such as the ability to infer referential intentions (Clark, 1990; Diesendruck & Markson, 2001). Clark (Clark, 1988) proposed that listener's inferences of speaker's referential intentions are guided by the Principle of Contrast: every two [linguistic] forms contrast in meaning. The principle is taken to entile that different names and descriptions refer to different objects. In later discussions, Clark's Principle of Contrast has been likened to Gricean cooperative maxims (de Marchena, Eigsti, Worek, Ono, & Snedeker, 2011; Diesendruck & Markson, 2001; Grice, 1975). For example, the use of a name that is unknown to the listener indicates that the speaker does not want to refer to the name-known object. If a speaker wanted to refer to an object whose name is mutually known, he will use that name, because speakers are expected to avoid ambiguity.

N3C (i.e. the novel name-nameless category principle) is a word learning principle that "novel names will be mapped onto unnamed categories" (Golinkoff et al., 1992; Mervis & Bertrand, 1995). According to Glolinkoff et al., adults and children are guided by N3C in disambiguation of novel names referents. Thus, they select the novel name-unknown objects since they belong to a category of which they don't know the name.

Finally, the *Feeling of Novelty* hypothesis proposed that infants and children expect novel names to map onto objects that feel novel (Merriman, Marazita, & Jarvis, 1995). The feeling of novelty elicited by an object depends on how recently the object, or another member of its category, has been encoded. Referent disambiguation based on this principle does not select a referent based on knowledge of its name. However, the principle has been proposed to account for experimental evidence that children and infants do novel name disambiguation. This was possible because in many experiments the name-unknown object was also a novel object, a member of an unknown category (Mather & Plunkett, 2012).

Halberda (Halberda, 2003, 2006) observed that the disambiguating inferences suggested by the Mutual Exclusivity Constraint or by the Pragmatic Accounts resemble a disjunctive syllogism. The name-unknown object is selected as referent of the novel name because the name-known object is rejected as referent (the exclusion is motivated by adherence either to a metalinguistic principle or to a pragmatic maxim). Such a disposition to infer an alternative by elimination of the others is in accordance with the rules of basic propositional logic. That the novel name does *not* denote the name-known object implies that it denotes the name-unknown object, granted that the novel name has been used to denote *either* one *or* the other object. Thus, attribution of basic logical representations offers a simple and elegant explanation of how the Mutual Exclusivity Constraint or the Principle of Contrast might be applied. Basic logical reasoning might be at the service of novel words disambiguation.

Novel name disambiguation has been documented in infants in their second year of life. It is therefore of primary importance for the purpose of this dissertation to attempt to assess to what degree novel name disambiguation is evidence of logical representation in preverbal infants.

1.3.2. Referent disambiguation and logical reasoning

Evidence of novel name disambiguation has been documented in infants in their second year of life. An increase in attention to name-unknown objects in response to a novel name has been documented in infants as young as 17 months (at 17 months (Halberda, 2003); at 20 months (Houston-Price et al., 2010). In a study by Mather and Plunkett (Mather & Plunkett, 2011) even 16-month-olds, who initially exhibited no increase in attention toward a name-unknown object in response to the novel name, were found to have mapped the novel name onto the name-unknown object by post-hoc comprehension tests. Moreover, infants as young as 12 months seem to infer the presence of a hidden object when a novel name is uttered in the absence of name-unknown objects (at 12 months (Fei Xu et al., 2005); at 15 months (Markman et al., 2003).

If novel name disambiguation is based on elementary logical reasoning, then the aforementioned studies might offer prima facie evidence of preverbal logical representation. However, there might be reasons to resist the antecedent of this conditional.

Explanations of novel name disambiguation based on either Mutual Exclusivity or the Pragmatic Accounts are suggestive of a disjunctive syllogism. However, as we have seen, alternative accounts have been proposed. Crucially, accounts based on N3C or Feeling of Novelty do not need to motivate the selection of the referent with the exclusion of its competitor(s). These accounts propose instead that the referent of a novel name is selected for its own properties, such as its membership of a name-unknown category or its (or its category's) novelty. But if novel name disambiguation is not motivated by the exclusion of the competitors, there is no need for the disjunctive syllogism.

Are there reasons that favor one explanation of novel name disambiguation against the others, especially when this phenomenon occurs in infants younger than 2 years?

Halberda (Halberda, 2006) reported eye-tracker results suggesting that adults and preschoolers (aged 3.8 months) exclude the name-known competitor before selecting the name-unknown object as a referent. However, to our knowledge, similar looking patterns have never been reported in infants or younger children.

Moreover, some of the few findings that are clearly indicative of novel name disambiguation during the second year of life (Halberda, 2003; Houston-Price et al., 2010; Mather & Plunkett, 2011) are compatible with all of the four accounts that we are contrasting. This is so because they adopt procedures that neither differentiate between direct referent selection and referent selection by exclusion, nor between disambiguation based on knowledge of an object's name and disambiguation based on stimuli novelty (since the the name-unknown object is a unfamiliar object (de Marchena, Eigsti, Worek, Ono, & Snedeker, 2011).

Other results in infants and toddlers apparently point to conflicting answers.

In a recent study, Mather and Plunkett (Mather & Plunkett, 2012) found evidence that object novelty might be sufficient for novel name disambiguation when word knowledge is not sufficient to determine the referent. Twenty-two-month-olds presented with two name-unknown objects, one of which they have been previously familiarized to, increased their looking at the less familiar object in response to a novel name. Mutual Exclusivity, the Pragmatic Accounts and N3C are grounded on name knowledge rather than object novelty. Thus, this result suggests that the Feeling of Novelty hypothesis might describe a mechanism that, at least starting from 22 months of age, supports novel name disambiguation.

An older study by Markman et al. (Markman et al., 2003) seems to point in the opposite direction. Presented with only one visible object, of which they know the name, infants as young as 15 months reacted to a novel name by either searching inside a box or searching around the room. This behavior suggests that infants expect the novel word to refer to a novel name-unknown object, even when no novel name-unknown object is visible. It is not clear, however, how the Feeling of Novelty hypothesis can account for such a finding: infants seem to do novel referent disambiguation when no object is present that exhibits higher novelty than the visible name-known competitor.

Finally, in a study by Fei Xu et al. (Fei Xu et al., 2005), 12-month-old infants seem to expect that more than one object is hidden inside a box after hearing a speaker call two distinct novel names while looking at the box. The development of such an expectation seems to point to some kind of novel name disambiguation. Inferring the presence of a second object seems to require the assumption that the second novel name cannot refer to the same object as the first novel name. Also, such a result conflicts with the Feeling of Novelty hypothesis, since no novel object, for which the novelty can be felt, is present.

While the findings of Markman et al. and Fei Xu et al. are in conflict with the Feeling of Novelty hypothesis, they can be elegantly accounted for by Mutual Exclusivity and the Pragmatic Accounts. The expectation that the referent of a novel name is present although not visible might be motivated by the rejection of the name-known object as a referent (and the plausible expectation that either a name refers to an object in sight or to an object out of sight).

An evaluation of the N3C principle is instead more complicated, since the principle is compatible with different accounts of its application. On the one hand, the N3C principle can be applied without rejection of the competitor, by direct selection of the object that belongs to a no name-known category. In this case it is not clear how the N3C principle can account for the aforementioned results, since the viewer has no direct evidence of the object and its category. On the other hand, the N3C principle can also be applied by rejecting the name-known object as a referent. In this case also N3C can be used to explain the expectation that the referent of the novel name is present although not visible. However, this type of application of N3C is suggestive of a disjunctive syllogism for exactly the same reason as Mutual Exclusivity and the Pragmatic Accounts (Halberda, 2006).

In summary, although novel name disambiguation has been documented in toddlers in their second year of life, and in one case in 12-month-old infants, it is not clear what strategies are behind those early episodes of disambiguation. The available evidence apparently conflicts, speaking in favor of different strategies (Feeling of Novelty hypothesis in Mather and Plunkett's study, Mutual Exclusivity and the Pragmatic Accounts in Marksman et al.'s and Xu et al.'s studies).

We formulated a tentative suggestion that might account for the apparently contrasting evidence reported. It is possible that multiple non-mutually-exclusive and individually-sufficient mechanisms of novel name disambiguation are available in the second year of life. Mather and Plunkett's findings suggest that cues of novelty might be sufficient for disambiguation by direct referent selection, when object name knowledge was insufficient. Studies by Marksman et al. and Xu et al. suggest that, when no cues of novelty are available, object name knowledge is used to infer the referent of a novel name, possibly by competitor rejection.

If this proposal is right, and it is not clear that this is so, then toddlers and, possibly, 12-month-old infants, under certain circumstances might do novel name disambiguation by competitor rejection. Thus, the results of studies by Marksman et al. and Xu et al. might be considered as potential clues of a disjunctive syllogism performed in word learning.

1.3.3. Invisible displacement task

Tasks designed to test novel word disambiguation tap into linguistic capacities. In such tasks, participants have to process stimuli that are either natural language expressions or artifact copies of them. Part of the ability required to solve the task is that used in speech perception and word learning.

Clarifying the relation between the development of logical representations and language acquisition might not be possible, if logical abilities are uniquely tested with tasks that are based on language processing. Hence, we judge it of primary importance to try to develop non-verbal tests of logical reasoning.

Steps have been made in this direction in comparative psychology, where non-verbal tasks designed to test basic logical inferences have been presented to human children, for the purpose of comparing them with animals of other species (Hill, Collier-Baker, & Suddendorf, 2012). The 164target logical inference was the disjunctive syllogism, often tagged as inferential reasoning by exclusion (Call &

Carpenter, 2000; Hill et al., 2012). The prototypical model of these tasks is the Piagetian invisible displacement task (Piaget, 1953). One object is hidden in one of two visually-occluded locations, but it is not known behind exactly which location the object is hidden. Then, one of the locations is revealed to be empty and the participant has to retrieve the hidden object. The correct solution is to search for the object in the location that has not been shown to be empty. Such an answer can be motivated by a disjunctive syllogism since the object is *either* in location A *or* in location B but the object is *not* in location B (since it has been shown to be empty), it entails by disjunctive syllogism that the object is in location A.

Studies based on the invisible displacement task have found that preschoolers succeed in selecting the location where the object was hidden (at 4 years (Premack & Premack, 1994); at 4 to 6 years (Watson et al., 2001); at 3 to 5 years (Hill et al., 2012). However, it is not clear if the children tested in those studies used a logical inference to infer the location of the object (Hill et al., 2012). A criticism suggests that the selection of the unexplored location in response to evidence of the emptiness of the other location might be the result of associative learning (Penn & Povinelli, 2007). For example, the children may have learned (and not inferred by logic) that when an object is hidden in a space with multiple hiding locations and all but one of those locations is found to be empty, usually the object is in the remaining location. Moreover, the relevant association does not even have to be learned during the test trials. The warm up trial may be sufficient. Alternatively, the association might instead be the result of past experience.

This type of criticism is addressed in a study by Watson et al. (Watson et al., 2001). Watson et al. adopted a version of the invisible displacement task with three hidden locations. No object was actually hidden in any of the locations, thus every time that a participant checked a hidden location, no object was found. Participants were left free to exhaustively check all locations one by one. The experimenters asked whether the motivation to check the next location, after having found the previous one empty, was grounded in associative learning or in a logical inference. The logic of the design was that, if participants' motivation was grounded in associative learning, finding two locations empty amounts to an extinction trial. Hence, it should slow down participants' searching rate. On the contrary, if participants' motivation was driven by disjunctive syllogism, finding two locations empty amounts to reaching certainty (or at least a higher degree of confidence) that the object is hidden in the remaining location. Thus it should result in a higher searching rate. A group of 4- to 6-year-old children and a group of domestic dogs were tested with this design. Watson et al. found that the dogs' speed of search decreased when they found the second empty location, suggesting that they relied on an associative strategy. In contrast, the children's speed of search did not decrease. The authors proposed that the speed of search of the children was indicative of a disjunctive syllogism.

Watson et al.'s findings suggest that children's performance is not based on association, since it is not simple to account for the difference of the responses of children and dogs in terms of an associative generalization. However, both Watson et al.'s design and the other invisible displacement task mentioned above might have failed to control for another confound. According to the account based on logical reasoning, the exclusion of one location causes a participant to search in the other *because* the exclusion of one location increases (by logic) the confidence that the object is in the other location (possibly to the level of certainty). However, there is an alternative explanation of why finding an empty location causes a participant to search in another. The alternative possibility is that the exclusion of one location does not minimally increase the participant's confidence that the object is in the other location. But, the participant checks the location that is still unchecked, because he is looking for the object and it is the only location that is left to be checked. In this case, checking the remaining location is not evidence of a disjunctive syllogism, since participant expectation regarding the last location has not been affected by finding the other location empty.

This type of explanation possibly applies in general to the Invisible Displacement tasks. Moreover, it is interesting to note that children in Watson et al.'s experiment do not increase their searching speed after finding two locations empty, which might be considered as evidence that their confidence regarding the presence of the object behind the third location has not increased, for the very same argument proposed by Watson et al. That is, it might be taken as evidence that they did not draw the conclusion that the object is in the third location. Or maybe they did, but this has not affected their search speed. Although Watson et al.'s study offers some evidence against associative learning, it does not offer conclusive evidence that the children used the disjunctive syllogism.

In summary, so far success in the invisible displacement task has not been documented in children younger than 2.5 years. Even in those cases, it is far from being clear that success in such tasks is based on the disjunctive syllogism. The strongest evidence of the logical inference in a non-verbal task is, in our opinion, offered by Watson et al. (Watson et al., 2001) with preschoolers aged from 4 to 6 years.

However, the interpretation of that result in terms of logical reasoning is not immune to criticism.

1.4. Goals

The first goal of this dissertation is to develop an experimental design and methodology apt to test the presence of logical abilities before the mastery of logical words, thus creating effective tools to begin the exploration of the origin and development of logical concepts in the first two years of life.

The second goal of this dissertation is to contribute to clarifying the relation between preverbal logical reasoning and early language processing and acquisition.

We have seen that the ability to disambiguate the referent of novel names, attested in the second year of life, might be linked to the logical rule disjunctive syllogism. Novel name referent disambiguation has been quite robustly documented in toddlers between 17 and 24 months (Halberda, 2003; Houston-Price et al., 2010; Mather & Plunkett, 2011) and in younger infants in two circumstances (Markman et al., 2003; Fei Xu et al., 2005). Thus, the aforementioned findings might be taken as evidence of preverbal logical reasoning. However, alternative accounts of novel name disambiguation have been proposed that do not rest on a disjunctive syllogism. Although the nature of the mechanisms behind novel name disambiguation is still unclear, we think that the presence of such ability in the second year of life can be seen as at least a preliminary clue of preverbal logical reasoning in the service of language processing and word learning.

In contrast, no evidence of logical inferences in experiments with non-linguistic stimuli have so far been reported in human infants younger than 2.5 years (but see Téglás' doctoral dissertation (Téglás, 2010) for a promising research on preverbal logical quantification). Interestingly, this asymmetry is compatible with a privilege relation of preverbal logical reasoning, and language learning and processing. That is, it is possible that before the third year of life, logical concepts are uniquely in the service of language processing or learning. However, the absence of evidence of logical inferences at the preverbal stage might simply reflect the need for more sensitive tasks.

As a result of these considerations, we think that a clarification of the relationship between language and logic in the early years of life demands the development of adequate non-linguistic procedures.

In the series of experiments presented in Chapter 2, we tested 12-month-olds and 19-month-olds with the Scooped Object task. The Scooped Object task is designed to test the ability to execute a disjunctive syllogism, without using any linguistic stimuli. In the task, participants are presented with animated movies and their cumulative looking time is measured and compared according to the violation of expectancy paradigm. We selected participants of either 19 months or 12 months of age. The first is an age at which novel names referent disambiguation is attested quite robustly (Halberda, 2003; Houston-Price et al., 2010; Mather & Plunkett, 2011), while the second is the earliest age at which that ability has ever been attested (Fei Xu et al., 2005). Across these experiments we developed a task apt to fairly test 12-month-old infants.

In the series of experiments presented in Chapter 3, we tested 12-month-olds and 19-month-olds with the Hidden Object task. The Hidden Object task is a task complementary to the Scooped Object task, designed to control for alternative, non-logic-based explanations of participants' performance.

In Chapters 4 we presented an analysis of participants' eye-movements, as recorded by an eye-tracker, in response to the Scooped Object task and the Hidden Object task. We analyzed (separately) both the eye-movements that accompanied participants' on-line processing of the movies and the eye-movements executed after the conclusion of the movie, during the measurement of participants' cumulative looking time.

2. EXPERIMENTS 1 TO 4: THE SCOOPED OBJECT TASK

2.1. Introduction

The Scooped Object task is designed to test the ability to perform a disjunctive syllogism without using linguistic stimuli and, hence, can be used in a population with limited linguistic capacity, such as infants and toddlers. A logical problem is presented in the form of animated cartoon movies. The movies show a scene with two objects, Object A and Object B, and a cup. The two objects have different shapes, textures and colors and seem to belong to different categories. However the upper part of the two objects is identical in any visual property (see Figure 1).

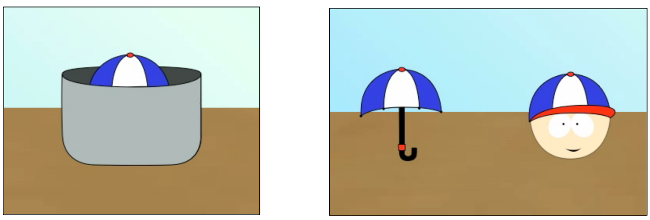

Figure 1. The ambiguous containment. When the lower part of the objects is covered, the two objects look identical.

The cup flies behind the occluder, reemerges from the center of its upper edge with one object inside, and finally lands in its original position. The visible upper part of the object inside the cup is that shared by the two objects.

The exit of the cup from behind the occluder with an object inside produces *the ambiguous containment*: Spatiotemporal and featural information are insufficient to determine which object is inside the cup. The spatiotemporal evidence is ambiguous. The trajectory of the cup is compatible with either one of the objects being moved by it. But one of the two objects has to be inside the cup, since no other object is present in the scene. The visible features of the object inside the cup do not help to disambiguate the spatiotemporal evidence. Therefore, there are two epistemic alternatives that are compatible with the ambiguous containment and one of them has to be correct. Inside the cup is *either* Object A *or* Object B.

The occluder is then removed, revealing the one object that has remained behind it, Object A. Afterwards, the retrieved object exits the scene, leaving the cup and its content. Finally, the cup is removed, revealing its content. Participants are tested with two types of conclusion. In one type of conclusion the object retrieved inside the cup looks identical to Object B. This is the Consistent Outcome. In the other type of conclusion the object retrieved inside the cup looks identical to Object A. This is the Inconsistent Outcome (see section n.m. for a detailed description of the movies).

An adult presented with such a sequence would easily realize that, since the cup had taken one of the two objects and Object A has been retrieved behind the screen, by elimination of disjunction, Object B is inside the cup. In other words, it is possible to infer what object is inside the cup by performing a disjunctive syllogism. For example, someone may reason through the following steps: "*either* the face is inside the cup *or* the umbrella is inside the cup", "the umbrella is *not* inside the cup" (since it is there in front of me), "so, the face is inside the cup".

If a participant presented with the Scooped Object task infers by elimination of disjunction which object is inside the cup, her expectation will be a *function of the object retrieved behind the occluder*. If

Object A is retrieved behind the occluder, then Object B has to be inside the cup. As a result, the participant will find the Inconsistent Outcome incongruent, where the object retrieved inside the cup looks like Object A. In contrast, she will find the Consistent Outcome acceptable, where the object retrieved inside the cup looks like Object B. We will refer to the hypothesis that participants presented with the Scooped Object task infer the content of the cup by disjunctive syllogism as the Disjunctive Syllogism hypothesis (DS hypothesis from this point on).

Studies on word learning have discovered a word learning strategy, often referred to as Mutual Exclusivity, that might be indicative of the ability to perform a disjunctive syllogism (see section n.m. for a discussion of the development of the Mutual Exclusivity strategy). Behavioral patterns suggestive of the Mutual Exclusivity strategy have been documented in young toddlers of 17 months of age (Halberda, 2003). This is one of the few published positive results that is potentially suggestive of logical inference in young toddlers or infants. Following this cue from the word learning literature, in Experiment 1 we tested whether 19-month-old toddlers are able to solve the Scooped Object task. We followed the violation-of-expectation method (i.e. unexpected events elicit longer looking time than expected ones (Baillargeon, Spelke, & Wasserman, 1985) and predicted that, if the toddlers infer by disjunctive syllogism what object is inside the cup, they will look at the Inconsistent Outcome for longer than the Consistent Outcome.

2.2. Experiment 1

2.2.1. Materials and methods

2.2.1.1. Participants

Twenty-two healthy full-term 19-month-old toddlers were retained for the analyses (8 girls, 14 boys, mean age: 19;02, range: 18;12-19;23). An additional 21 toddlers were tested but not retained due to either fussiness (9), caretakers' interaction (3), equipment failure (2), experimenter error (6), or participant unresponsiveness (1). The participants' parents were contacted by telephone and were given a small present and a certificate of attendance.

2.2.1.2. Materials

The stimuli of Experiment 1 consisted of 28 animation movies (4 familiarization movies and 24 test movies) representing objects that interact with an occluder and/or a container. The movies were initially prepared as a series of animated slides with the software Keynote 5.0 (apple iWork'09 package). The slides were conceived as modules that, when joined in different combinations, generate the 28 movies. The slides were exported as QuickTime movies generated at 60 fps, 1024 × 768 pixels, MPEG-4 Video compression. The movies were played on a 24-inch screen with the software PsyScope X (http://psy.ck.sissa.it/) running on an Apple Mac Pro Quad Core 2.8 computer. The screen covered a 35 × 26.5-cm area. The movies were accompanied with sounds that were contingent with the events taking place in the movies. The sounds were associated with the elements of the scene, for example: the occluder rising and falling was associated with a specific sound, the cup's movements were associated with another specific sound, and so on.

2.2.1.2.1. Familiarization movies

In Experiment 1 we used four movies to familiarize the participants with the animated representation of the events of occlusion and containment as well as with the objects that were later used in the test (see APPENDIX1 for pictures and detailed descriptions of the content of the familiarization movies).

The familiarization movie F1a was designed to introduce the function of the occluder. In the movie, two objects (a white bear and a red car) are initially covered by an occluder and then retrieved once the occluder is removed (see APPENDIX1, section 1).

The Familiarization movies F2a and F3a were designed to introduce the functions of the cup and the occluder when the two "act" separately (i.e. the cup scoops the object only after the removal of the occluder). In the movie, one object (a white bear in F2a and a red car in F3a) is initially covered by an occluder and retrieved once the occluder is removed. Afterwards the cup scoops the object, carries it to a different location and eventually releases it (see APPENDIX1, section 2).

The movie F4a was designed to introduce all test objects. The six objects lie side by side and move one after the other to attract the viewer's attention (see APPENDIX1, section 3).

Crucially, although during the familiarization movies participants were presented with the objects and the events used in the test, the solution of the Scooped Object task was never presented.

That is, during familiarization, participants never saw which object was inside the cup after that the cup has scooped one of the two objects behind the occluder and one of the two objects had been retrieved outside the occluder.

2.2.1.2.2. Test movies

For Experiment 1 we prepared 24 test movies. Across the 24 test movies, we used three different object pairs (see Figure 2).

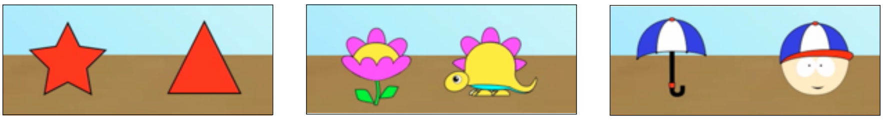

Figure 2. Object pairs of Experiment 1. The Star/Triangle pair consists of a red star and a red triangle of the same approximate size. The Flower/Dinosaur pair consists of a yellow, pink and green flower and a yellow, pink and cyan dinosaur.

The objects in this pair have different shapes, textures, colors and present cues of the category contrast AGENT and NON-AGENT (Surian & Caldi, 2010). The Umbrella/Face pair consists of one pink human head with wide white eyes and a blue, red and white cap, and a blue and white umbrella with a black handle with red details. The objects in this pair have different shapes, textures, colors and present cues of the category contrast HUMAN and NON-HUMAN (L. L. Bonatti, Frot, Zangl, & Mehler, 2002; L. L. Bonatti, Frot, & Mehler, 2005). These objects are different from one another not just in visual features and category membership, but also in the characteristic movements hey perform when they are introduced at the beginning of the scene. The face smiles and then tilts slightly to its left and back. The dinosaur jumps and nods its head up and down a pair of times.

The umbrella and the flower oscillate left and right several times. The triangle and the star pulsate repeatedly. However, and crucially for the purpose of the experiment, the objects in each pair have an identical upper part, so they look identical if partially hidden inside the cup.

Each test movie presents one of the two outcomes of the Scooped Object task (see Figure 3).

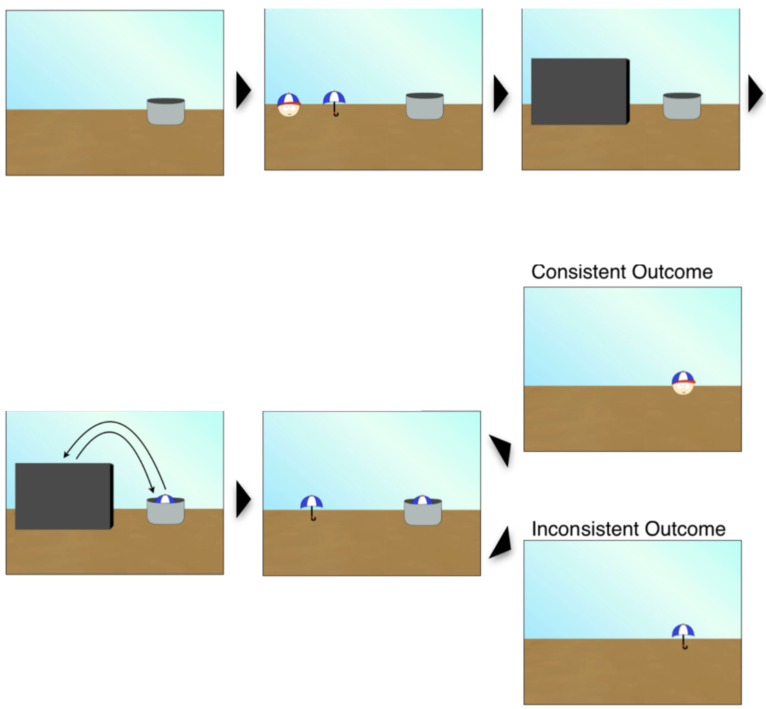

Figure 3. Schema of Scooped Object task movies. In half of the test movies, the object revealed by the cup looks identical to the other object in the pair (Consistent Outcome condition). In the other half of the test movies, the object revealed by the cup looks identical to the object that has just left the scene (Inconsistent Outcome condition).

A test movie begins with an empty grey cup lying in the right region of the scene. After approximately 0.50 s, two objects enter the scene, one after the other, by falling from the upper side of the movie frame.

The objects land side by side in the left side of the scene. Immediately after the second object has landed on the right side of the first object, the two objects sequentially execute their characteristic movements, that is: the first object moves for 2 s while the second is motionless then, as soon as the first object has stopped moving, the second object starts to move for 2 s while the first is motionless. Each characteristic movement is accompanied by a sound. After the characteristic movement, a dark grey occluder emerges from the portion of ground below the two objects and rises up to completely cover the two objects. As soon as the occluder has completely risen, the cup flies behind the occluder. The cup enters the occluder through the exact center of its upper edge. After, the cup reemerges from the occluder with one object inside, and lands in the right region of the scene. Now the cup is motionless and the upper part of the object is visible from inside the cup. From inside the cup the object pulses for 2 s while a cuckoo sound is played. The occluder then falls and disappears below the ground, revealing one of the two objects. The object is now occupying the center of the region that was covered by the occluder. Once revealed, the object that was covered by the occluder executes its specific characteristic movement, accompanied by the cuckoo sound. Immediately after the visible object has finished moving, the object partially occluded inside the cup pulses, accompanied by the cuckoo sound. The visible object and the partially occluded object remain motionless for about 1 s and then the visible object exits the movie frame via the left side. After the visible object has exited the scene, the cup reveals its content by falling and disappearing under the ground. In half of the test movies, the object revealed by the cup looks identical to the other object in the pair. This is the Consistent Outcome. In the other half of the test movies, the object revealed by the cup looks identical to the object that has just left the scene. This is the Inconsistent Outcome. The total length of each test movie is 20.35 s.

Across the 24 test movies, the following variables were covered:

1. Outcome Type. Two levels: Consistent or Inconsistent.
2. Pair. Three levels: Star/Triangle or Flower/Dinosaur or Umbrella/Face.
3. Object; the type of the object eventually released by the cup. Six levels: Star, Triangle, Flower, Dino, Umbrella or Face; nested in Pair.
4. Distance; the distance, at the beginning of the movie, between the cup and the object eventually released by the cup. Two levels: Distant or Close (see Figure 4).

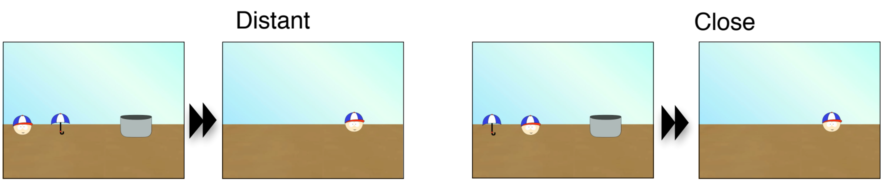

Figure 4. The variable Distance. In half of the test movies, the object eventually released by the cup is the object that, at the beginning of the movie was the one less close to the cup (Distant condition). In the other half of the test movies, the object eventually released by the cup is the object that, at the beginning of the movie was the one closer to the cup (Close condition).

Within each pair, the position of the objects at the beginning of the movie was inverted in half of the movies.

2.2.1.3. Procedure

The experiment took place in a sound-attenuated room with dimmed lighting. During the full experimental period, participants were seated on their parent's laps at about 60 cm distance from the stimuli presentation monitor. The parents wore opaque glasses that prevented them from seeing the stimuli. Parents were instructed to keep the child seated on their lap and to not interact with them. The experimenter sat outside the room, controlling the experiment and monitoring participants' behavior from a screen. The experimenter was blind to the experimental condition since the experimenter screen does not present a view of the stimuli, but just a frontal shot of the participants taken by an infrared camera inside the experimental room. The presentation of the stimuli was participant-controlled. That is, during the stimuli presentation, the movie was paused from when the toddler was not looking at the screen until they reoriented toward the screen. This procedure ensures that each toddler sees the entirety of each movie.

The experiment consisted of a Familiarization Phase and a Test Phase. In the Familiarization Phase of the current experiment participants saw four movies. Participants began by viewing a movie that introduced the occluder function. In the second and third familiarization trials they saw two movies that introduced the cup and occluder functions separately. Eventually, they saw a movie that introduced all test objects (for a summary of the familiarization structure see Table 1).

Trial	Movie	Function	Length	Reference
1	F1a	familiarization with the occluder	1 x 12.31 s	APPENDIX1 section 1
2-3	F2a-F3a counterbalanced order	familiarization with cup and occluder	2 x 15.81 s	APPENDIX1 section 2
4	F4a	familiarization with the test objects	1 x 13 s	APPENDIX1 section 3

Table 1. Summary of the structure, functions and length of the familiarization of Experiment 1.

In the Test Phase, participants saw six test movies. The movies were administered as follows:

1. Half of the participants were presented with Consistent or Inconsistent Outcomes in the sequence CIICCI, while the other half where presented with the sequence ICCIIC ("I" stands for "Inconsistent Outcome" and "C" stands for "Consistent Outcome".
2. Similarly, the Distance condition was administered either in the sequence CDDCCD or in the sequence DCCDDC ("C" stands for "Close" and "D" stands for "Distant").
3. Each participant was tested two times with each object pair. The order of presentation of the pairs was either $C_1C_2C_3C_1C_2C_3$ or $C_2C_3C_1C_2C_3C_1$ or $C_3C_1C_2C_3C_1C_2$ ("C_1" stands for one pair, "C_2" stands for the second pair, "C_3" stands for the remaining pair).
4. We planned to present each participant with six trials, each trial terminating with a distinct object released by the cup. However, because of a mistake in planning the administration of the movies, for each participant there was exactly one repetition. In one of the

six trials, the object inside the cup was repeated. In other words, of the six possible objects, 4 objects were released by the cup in exactly one trial each, 1 object was repeated in two trials and 1 object was not used in any trials. Luckily, despite the mistake, this variable was counterbalanced across participants, since which object was never released by the cup was equally repeated across participants.

At the end of each movie, the last frame remained paused on the monitor and looking time monitoring began. The experimenter monitored looking time by pressing the mouse when a participant was directing their gaze towards the screen, and releasing it when the participant looked away. A trial lasted when participants looked away for at least 2.5 s or when they looked at the screen for a total of 35 s. These criteria were selected to allow for a 10-15% possible experimental error when tracking time online. Data were then coded offline frame by frame, with the software Psycode (http://psy.ck.sissa.it/PsyCode/PsyCode.html). For offline coding, a trial was considered complete when the participant looked away for 2 s or when the cumulative looking time exceeded 30 s.

2.2.2. Results

Participants were excluded from the analyses if they had less than one valid trial per Outcome Type (Consistent or Inconsistent) or if they had a cumulative looking time of 30 s in more than half of the trials. A trial was considered invalid, and not included in the analyses, if either the caretaker interacted (verbally or not) with the participants differently from the received instructions during the measurement phase, the experimenter erroneously triggered the end of a trial before a 2 s look-away period, the participant looked at the outcome for less than 1 s, or the looking time exceed 2.5 absolute median deviations from the median (Leys, Ley, Klein, Bernard, & Licata, 2013), computed per condition. The trial rejection rate due to the filter based on the deviation from the median was 12.5% of the Inconsistent Outcome trials and 16.1% of the Consistent Outcome trials.

Preliminary analyses (see APPENDIX2) detected no interaction effects of Outcome Type with Pair (Star/Triangle vs Dinosaur/Flower vs Face/Umbrella) or Distance (Close vs Distant). Therefore the data were collapsed across Pair and Distance in the subsequent analysis.

For each participant, two mean looking time scores were calculated by averaging looking time trial scores by Outcome Type (Consistent or Inconsistent). To test for a difference in performance between the two outcome types, an analysis of variance (ANOVA) with Outcome Type as a within-participants factor and Participant as random variable was performed on mean looking time scores. The analyses detected a main effect of the Outcome Type ($M_{Consistent}$ = 7.1 s, $M_{Inconsistent}$ = 10.7 s; $F(1, 21) = 9.29$, $P < 0.01$), showing that the 19-month-olds looked at the Inconsistent Outcome for longer than the Consistent Outcome. (see Figure 5). A Wilcoxon Signed-ranks test confirmed the main effect of the Outcome Type ($z = -2.48$, $P < 0.05$, two-tailed).

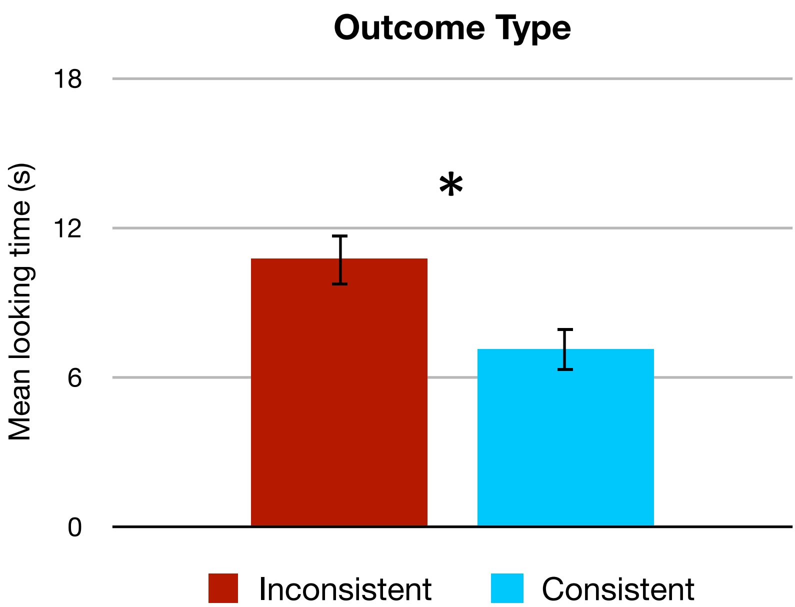

Figure 5. Nineteen-month-olds' mean looking times at the Inconsistent Outcome and Consistent Outcome in Experiment 1. Error bars represent SEs.

2.2.3. Discussion

The 19-month-old toddlers looked at the Inconsistent Outcome for longer than the Consistent Outcome, a result predicted by the DS hypothesis.

Thus, toddlers' performance makes a prima facie case that the representational resources required to perform a disjunctive syllogism may already be available at the age of 19 months.

As we have seen (sections 1.3.2.), nineteen months is an age at which novel names referent disambiguation is attested quite robustly (Halberda, 2003; Houston-Price et al., 2010; Mather & Plunkett, 2011). However, in two circumstances such ability has been documented in younger infants (Markman et al., 2003; Fei Xu et al., 2005). In the next experiment we presented 12-month-old infants with the Scooped Object task, infants of the earliest age at which novel names referent disambiguation has ever been attested.

2.3. Experiment 2

2.3.1. Materials and methods

Materials and methods of Experiment 2 were identical to those of Experiment 1.

2.3.1.1. Participants

Twenty-four healthy full-term 12-month-old infants were retained for the analyses (14 girls, 10 boys, mean age: 12;06, range: 11;21-12;14). An additional 36 infants were tested but not retained due either to fussiness (22), caretakers' interaction (2), equipment failure (3), experimenter error (8), or medical history (1). The participants' parents were contacted by telephone and were given a small present and a certificate of attendance.

2.3.1.2. Procedure

The procedure of Experiment 2 was identical to the one of Experiment 1.

2.3.2 Results

Rejection criteria were identical to those of Experiment 1. The trial rejection rate due to the filter based on the deviation from the median was 10.3% of the Inconsistent Outcome trials, and 10.3% of the Consistent Outcome trials.

Preliminary analyses (see APPENDIX2 for the complete results of the analyses) detected no interaction effects of Outcome Type with Pair (Star/Triangle vs Dinosaur/Flower vs Face/Umbrella). A two-way ANOVA with Outcome Type and Distance (Close vs Distant) as between-participants factors and Participant as a random variable detected a trend toward an interaction effect ($M_{Close*Inconsistent}$ = 9.4 s, $M_{Close*Consistent}$ = 6.8 s, $M_{Distant*Inconsistent}$ = 7.5 s, $M_{Distant*Consistent}$ = 8.4 s; $F(1, 44) = 3.96$, $P = 0.053$). Post hoc Scheffé tests revealed that in the Close condition there was a trend toward a difference between looking time at the Inconsistent Outcome and looking time at the Consistent Outcome ($P < 0.083$). All the other pairwise comparisons did not detect a significant difference. Either such a tendency toward an interaction effect was not detected in other experiments (Experiment 1, Experiment 4 Experiment 5, Experiment 6) or an effect in the Distant condition (and not in the Close) was detected in Experiment 3, suggesting that the effect may be due to group variability.

For each participant, two mean looking time scores were calculated by averaging looking time trial scores by Outcome Type (Consistent or Inconsistent). To test for a difference in performance between the two outcome types, an ANOVA with Outcome Type as a within-participants factor and Participant as random variable was performed on mean looking time scores. The analyses detected no effect of the Outcome Type ($M_{Consistent}$ = 8.2 s, $M_{Inconsistent}$= 8.7 s; $F(1, 23) = 0.499$, $P = 0.5$), showing that infants' looking time at the two outcome types was about the same. (see Figure 6).

A Wilcoxon Signed-ranks test detected no main effect of Outcome Type (z = -1.05, $P = 0.29$, two-tailed).

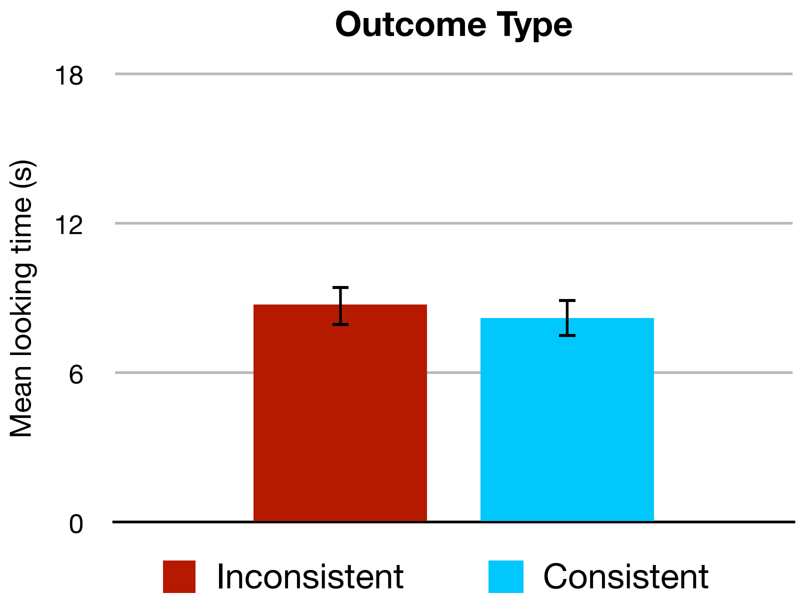

Figure 6. Twelve-month-olds' mean looking times at the Inconsistent Outcome and Consistent Outcome in Experiment 2. Error bars represent SEs.

Finally, an ANOVA with Outcome Type as a within-participants factor, Experiment as a between-participants factor and Participant as random variable was performed on mean looking time scores to compare the performance of the 12-month-old infants (Experiment2) with the performance of the 19-month-old toddlers (Experiment1). There was a main effect of Outcome Type ($M_{Consistent}$ = 7.7 s, $M_{Inconsistent}$ = 9.7 s; $F(1, 44) = 9.39$, $P < 0.005$) and no main effect of Age ($M_{Experiment2}$ = 8.5 s, $M_{Experiment1}$ = 8.9 s; $F(1, 44) = 0.15$, $P = 0.69$). Crucially, an interaction between Outcome Type and Age was detected ($F(1, 44) = 0.15$, $P < 0.05$). (see Figure 7). Post hoc Scheffé tests revealed that 19-month-olds' looking time at the Inconsistent

Outcome was higher than 19-month-olds' looking time at the Consistent Outcome ($P < 0.001$), higher than 12-month-olds' looking time at the Inconsistent Outcome ($P < 0.05$) and higher than 12-month-olds' looking time at the Consistent Outcome ($P < 0.05$). Twelve-month-olds' looking time at the Inconsistent Outcome and 19-month-olds' looking time at the Consistent Outcome in Experiment 2 were about the same ($P < 0.61$). All the other pairwise comparisons did not detect a difference. This pattern of results confirmed that toddlers reacted with a higher looking time at the Inconsistent Outcome, while infants' looking time at the Inconsistent Outcome and infants' or toddlers' looking time at the Consistent Outcome were about the same.

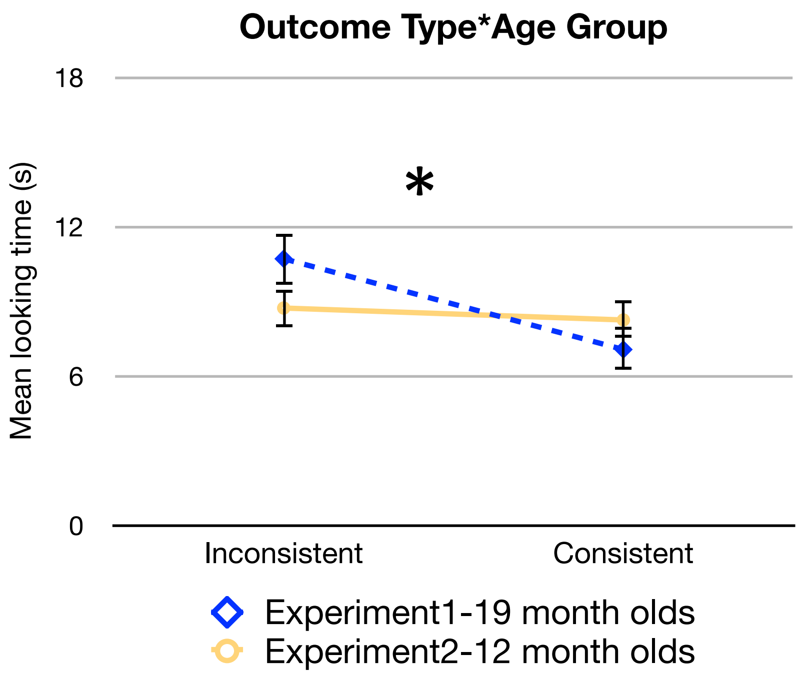

Figure 7. Mean looking times at the Inconsistent Outcome and Consistent Outcome in function of the Experiment (Experiments 1 and 2). Error bars represent SEs.

2.3.3. Discussion

The 12-month-old infants' looking time at the Inconsistent Outcome did not differ from their looking time at the Consistent Outcome. This result suggests that Infants did not infer by elimination of disjunction which object was inside the cup.

It is possible that the 12-month-old infants did not deploy the appropriate disjunctive syllogism, because at that age infants have not yet acquired the ability to execute such a logical inference.

A second alternative possibility is that at 12 months of age, the disjunctive syllogism is uniquely in service of language processing, and thus it can not be deployed to disambiguate the identity of an hidden object (see section 1.4.).

Yet, a third possibility is that 12-month-olds' failure is an artifact of our procedure. Possibly, the current task is too hard to fairly test disjunction elimination at that age. Some aspects of the task may be excessively demanding for younger infants in terms of non-logical resources. In this case, infants may have failed in the task even if they had the ability to perform a disjunctive syllogism in a non-linguistic task.

A first potential difficulty factor in our task is the length of the occlusion. In the current task the time elapsed from the moment that the two objects are entirely covered by the occluder up to the moment that the first object is retrieved from behind the occluder is about 7 s, and the time elapsed up to the moment that a second object is released by the cup is about 13.5 s. As explained in the discussion of the first potential difficulty factor, in order to detect the Inconsistent Outcome by a disjunctive syllogism one has to remember the presence of the two objects and their features (perhaps, especially their category memberships) up to the moment that those objects are retrieved. About 13 s is a rather long occlusion compared to the standard of task administered in the object individuation and identification literature. For example, in comparable procedures used by Tremoulet and colleagues with 12-month-old infants, the time elapsed from the last appearance of an occluded object up to its retrieval is about 4-5 s (Tremoulet, Leslie, & Hall, 2000). Plausibly, the longer the occlusion lasts, the higher the risk that a participant may forget some piece of information. It is also plausible that the effect of the length of the occlusion on information retention is modulated by age, with older infants outperforming younger ones. The length of the occlusion might have prevented 12-month-olds, but not 19-month-olds, from maintaining in memory the representations of the occluded objects or their features for as long as needed, and thus might have caused the infants' failure.

A second possible difficulty factor behind the 12-month-old infants' failure might be a lack of evidence clear enough to lead them to recognize the ambiguity of the containment event. One reason for representing the containment event in terms of a disjunction (e.g. *either* the triangle is inside the cup *or* the star is inside the cup) is the recognition of two epistemic alternatives: the possibility that the triangle is inside the cup and the possibility that the star is inside the cup. The recognition of these two epistemic alternatives might play a role in the formation of a disjunctive representation of the scene.

Two aspects of the task may cue the recognition of such epistemic alternatives: that the cup can scoop either one of the two objects and that the upper part of the two objects look identical. However, evidence that the cup can scoop either of the objects is never presented during the familiarization movies of Experiment 2, since in the movies that familiarized the participants with the scooping event only one object is present in the scene. Furthermore, the only opportunity to directly see the way a test object looks when partially occluded is in the test movies when the cup exits from behind the occluder with one of the two objects inside. As a result, the available evidence of the epistemic alternatives might have been insufficient for inducing the infants to think of the containment event disjunctively, while sufficient for the toddlers.

In summary, we have identified two potential difficulty factors that may have prevented 12-month-old infants from demonstrating their hypothetical ability to perform a disjunctive syllogism:

1. The length of the occlusion.
2. The poorness of the available evidence of the ambiguity of the containment event.

In Experiment 3 we tested whether, once these factors are reduced, 12-month-old infants can solve the Scooped Object task. Following the proposed diagnosis of the potential difficulty factors of the task, we prepared a slightly different version of the task used in Experiments 1 and 2. In the new version of the task, the nature of the logical problem was left unchanged but:

1. The occlusion length was reduced: the time elapsed from the moment that the two objects were entirely covered by the occluder up to the moment that the last of the two objects was released by the cup was about 5.2 s.
2. Additional evidence of the ambiguity of the containment event was provided both in the familiarization and the test movies.

2.4. Experiment 3

2.4.1. Materials and methods

Materials, methods and procedure of Experiment 3 were identical to Experiment 2 except for the changes specified in the following sections.

2.4.1.1. Participants

Twenty-three healthy full-term 12-month-old infants were retained for the analyses (15 girls, 8 boys, mean age: 12;07, range: 11;21-12;14). An additional 11 infants were tested but not retained due either to fussiness (10) or experimenter error (1). The participants' parents were contacted by telephone and were given a small present and a certificate of attendance.

2.4.1.2. Materials

2.4.1.2.1. Familiarization movies

In Experiment 3 we used 8 brand new familiarization movies (see APPENDIX1 for pictures and detailed descriptions of the content of the familiarization movies).

The familiarization movies F1b and F2b were designed to introduce the functions of the cup and of the occluder when the two "act" separately (i.e. the cup scoops the object only after the removal of the occluder). In each movie, an object (a red puppet in F1b and a red star in F2b) is covered by an occluder and then retrieved once the occluder is removed. Afterwards the cup takes the object, carries it to a different location and eventually releases it (see APPENDIX1, section 4).

The familiarization movies F3b, F4b, F5b and F6b were designed to introduce the interaction of the functions of the cup and the occluder (i.e. the cup scooped one object from behind the occluder and brought it out). It also presents to the infants the way each test object looks when partially hidden inside the cup. In each movie, an object (the flower in F3b, the dinosaur in F4b, the umbrella in F5b and the face in F6b) is covered by an occluder. The cup then scoops the object behind the occluder and carries it to a different location. When the occluder and the cup are removed, the object is released by the cup (see APPENDIX1, section 5).

The familiarization movies F7b and F8b were designed to introduce the interaction of the functions of the cup and the occluder, when two objects are hidden behind the occluder, without ever showing the outcome (what object is hidden in the cup). The presentation of the scooping event in familiarization may support the recognition of the ambiguity of the containment in the test. In each movie, two objects (the red puppet and the red star) are covered by an occluder. Afterwards the cup scoops an object behind the occluder and carries it to a different location. Once the occluder is removed, an object is found behind it (the red puppet in F7b and the red triangle in F8b) while the other object remains hidden inside the cup (see APPENDIX1, section 6).

2.4.1.2.2. Test movies

For Experiment 3 we prepared 16 test movies. The Test movies of Experiment 3 had an identical structure to the Test movies of Experiment 2 except for the following changes.

First, in order to increase participants' interest in the task we eliminated the movies presenting the Star/Triangle pair, since in Experiment 2 participants had a non-significant tendency to look at that pair less.

Secondly, after the objects' characteristic movement, the occluder did not immediately rise up to entirely cover the two objects. Instead it covered the lower half of the two objects, leaving visible only their upper half. That was the part of the two objects that looked identical, and the only visible part of the objects during the partial occlusion inside the cup (see Figure 8).

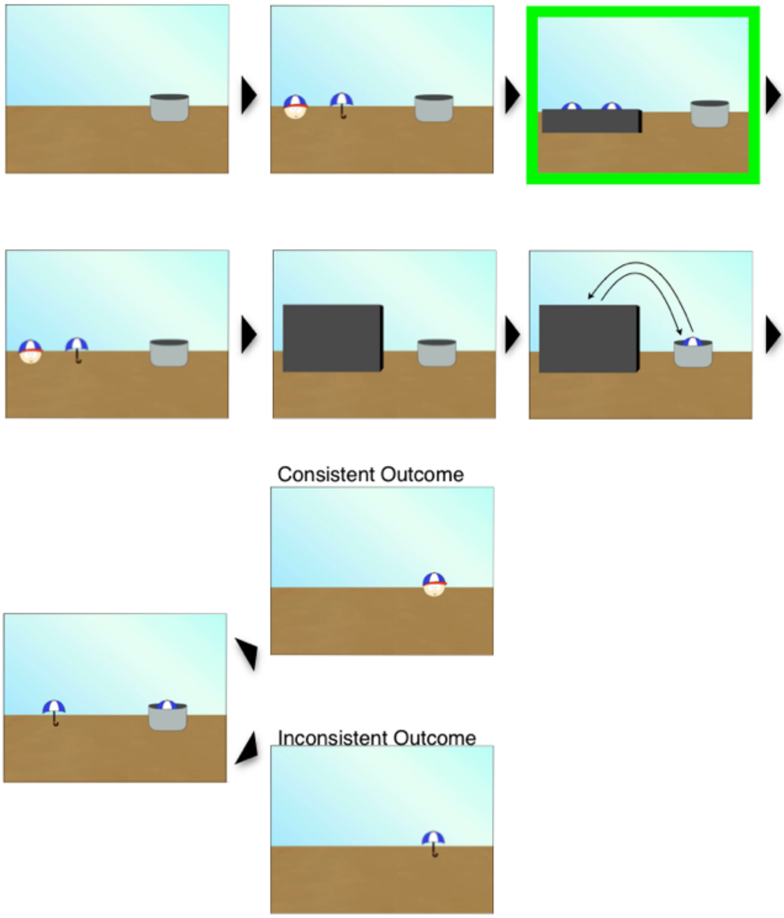

Figure 8. *The test movies of Experiment 3 offered additional evidence that the upper part of the two objects looked identical. In a new part of the test movie the occluder remained stationary in this state for 1 s and then it fell and disappeared under the ground, revealing both of the objects in the same state as before the occlusion. The objects remained entirely visible for about 1 s and then the occluder rose again to cover the lower half of the objects. The occluder remained stationary in this state for 1 s and eventually it rose to entirely cover the two objects. Afterwards, the movie continued exactly as in the Test movies of Experiment 2.*

This new part of the test movie was designed to mitigate one of the potential difficulty factors of Experiment 2 by offering some additional evidence of the ambiguity of the containment event. Specifically, this part of the test movie showed to the participants the way the members of the test pairs looked when partially occluded, and that the upper part of the two objects looked identical. A more direct demonstration of this crucial property of the test pairs might help with the recognition of the two epistemic alternatives compatible with the containment event.

Thirdly, the events in the test movies of Experiment 3 were generally faster than in Experiment 2 and the objects never executed any movement from inside the cup. These changes reduced the overall occlusion length (the time elapsed from the moment that the two objects are entirely covered by the occluder to the moment that the last of the two objects is released by the cup) to about 5.2 s. Thus, the memory demand of the task was also mitigated. The total length of each test movie was 16.63 s.

2.4.1.3. Procedure

In the Familiarization Phase of the current experiment participants saw 8 movies. Participants began by viewing two movies that separately introduced the cup's and occluder's functions. From the third to the sixth familiarization trials they saw four movies that introduced the interaction between the cup and the occluder. Finally, they saw two movies that introduced the test event structure, except for the equivalent of the outcome (for a summary of the familiarization structure see Table 2).

Trial	Movie	Function	Length	Reference
1-2	F1b-F2b counterbalanced order	familiarization with cup and occluder	2 x 12.73 s	APPENDIX1 section 4
3-6	F3b-F6b random order	familiarization with interaction of cup and occluder and with test objects.	4 x 10.38 s	APPENDIX1 section 5
7-8	F7b-F8b counterbalanced order	familiarization with test trial structure	2 x 17.98 s	APPENDIX1 section 6

Table 2. Summary of the structure, functions and length of the familiarization of Experiment 3.

The structure and the procedure of the Test Phase of the current experiment was identical to that of Experiment 1 except for two changes. First, the number of test trials was reduced to 4, since the movies presenting the Star/Triangle pair were eliminated. Secondly, the movies were administrated as follow:

1. Half of the participants were presented with Consistent or Inconsistent Outcomes in the sequence CIIC, while the other half were presented with the sequence ICCI ("I" stands for "Inconsistent Outcome" and "C" stands for "Consistent Outcome").
2. Similarly, the Distance condition was administered either in the sequence CDDC or in the sequence DCCD ("C" stands for "Close" and "D" stands for "Distant").
3. Each participant was tested two times with each object pair. The presentation order of the pair was either $C_1C_2C_1C_2$ or $C_2C_1C_2C_1$ ("C_1" stands for one pair, "C_2" stands for the other pair).

4. Each participant was presented with all four levels of the variable Object (the type of object retrieved inside the cup). Across participants, the trial number was crossed with which object was found inside the cup.

2.4.2. Results

Rejection criteria were identical to those of previous experiments. The trial rejection rate due to the filter based on the deviation from the median was 2.1% of the Inconsistent Outcome trials and 10.9% of the Consistent Outcome trials.

A two-way ANOVA with Outcome Type and Distance (Close vs. Distant) as between-participants factors and Participant as a random variable nested in both Outcome and Distance detected a main effect of Outcome Type ($M_{Consistent}$ = 8.3 s, $M_{Inconsistent}$ = 11.7 s, $F(1, 42) = 5.43$, $P < 0.05$), no effect of distance (M_{Close} = 9.6 s, $M_{Distant}$ = 10.6 s; $F(1, 42) = 1.08$, $P = 0.304$) and an interaction effect ($M_{Close*Inconsistent}$ = 9.5 s, $M_{Close*Consistent}$ = 9.6 s, $M_{Distant*Inconsistent}$ = 14 s, $M_{Distant*Consistent}$ = 7 s; $F(1, 42) = 5.64$, $P = 0.022$); (see Figure 9). Post hoc Scheffé tests revealed that in the Distant condition looking time at the Inconsistent Outcome was higher than looking time at the Consistent Outcome ($P < 0.005$), while in the Close condition looking time at the Inconsistent Outcome was not different than looking time at the Consistent Outcome ($P = 0.975$). Looking Time at the Inconsistent Outcome in the Distant condition was also higher than both looking time at the Inconsistent Outcome in the Close condition ($P < 0.05$) and looking time at the Consistent Outcome in the Close condition ($P < 0.05$). All other pairwise comparisons did not detect a significant difference. Either such an interaction effect between Distance and Outcome Type was not detected in other experiments (Experiment 1, Experiment 4, Experiment 5, Experiment 6) or else a tendency toward the opposite trend was detected in Experiment 2, suggesting that the effect may be due to group variability.

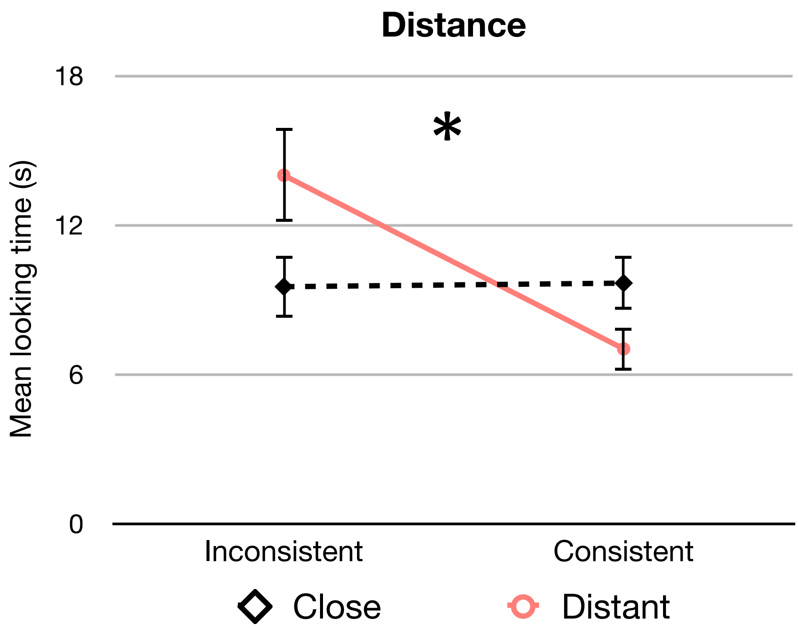

Figure 9. Mean looking times at the Inconsistent Outcome and Consistent Outcome in function of Distance (Experiment 3). Error bars represent SEs.

A two-way repeated measures ANOVA with Outcome Type and Pair as within-participants factors and Participant as random factor detected a main effect of Outcome Type ($F(1, 60) = 8.32$, $P < 0.01$). A main effect of Pair was also detected ($M_{Flower/Dinosaur} = 8.2$ s, $M_{Umbrella/Face} = 12$ s; $F(1, 60) = 7.66$, $P < 0.01$). Infants tended to look longer at the outcomes with the Umbrella/Face pair than at the outcomes with the Flower/Dinosaur pair. Crucially, an interaction effect between Outcome Type and Pair was found ($M_{Umbrella/Face*Inconsistent} = 14.6$ s,

$M_{Umbrella/Face*Consistent}$ = 9.1 s, $M_{Flower/Dinosaur*Inconsistent}$ = 8.9 s, $M_{Flower/Dinosaur*Consistent}$ = 7.5 s; $F(1, 60) = 4.43$, $P < 0.05$); (see Figure 10). Post hoc Scheffé tests revealed that with the Umbrella/Face pair, looking time at the Inconsistent Outcome was higher than looking time at the Consistent Outcome ($P < 0.001$), while with the Flower/Dinosaur pair, looking time at the Inconsistent Outcome did not differ from looking time at the Consistent Outcome ($P = 0.58$). Looking Time at the Inconsistent Outcome with the Umbrella/Face pair was also higher than both looking time at the Inconsistent Outcome with the Flower/Dinosaur pair ($P < 0.001$) and looking time at the Consistent Outcome with the Flower/Dinosaur pair ($P < 0.0005$). All other pairwise comparisons did not detect a significant difference. This pattern of looking times suggests that the Inconsistent Outcome caused infants to look longer only when it was instantiated by the Umbrella/Face pair.

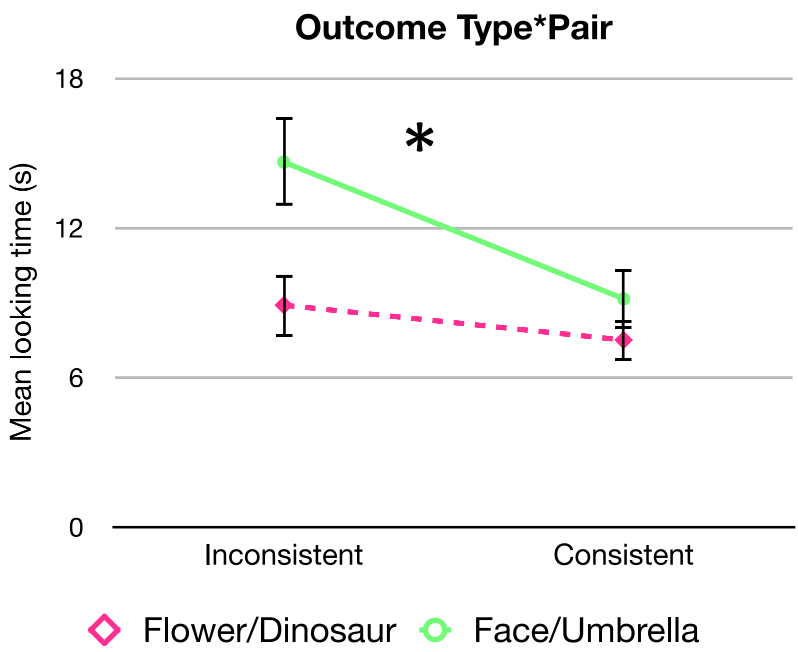

Figure 10. Mean looking times at the Inconsistent Outcome and Consistent Outcome in function of Pair (Experiment 3). Error bars represent SEs.

In order to compare the effect of Outcome Type on participants' looking times in Experiment 2 and Experiment 3, two mean looking time scores were calculated for each participant by averaging looking time trial scores by Outcome Type (Consistent or Inconsistent) and then performing a two-way repeated measures ANOVA with Outcome Type as a within-participants factor, Experiment as a between-participants factor and Participant as a random factor nested in Experiment. There was a main effect of Outcome Type ($M_{Consistent}$ = 8.33 s, $M_{Inconsistent}$ = 10.13 s; $F(1, 45) = 8.06$, $P<0.01$) and no effect of Experiment ($M_{Experiment2}$ = 8.5 s, $M_{Experiment3}$ = 10 s; $F(1, 45) = 1.9$, $P = 0.17$). Crucially, an interaction between Outcome Type and Experiment was detected ($M_{Experiment2*Consistent}$ = 8.24 s, $M_{Experiment2*Inconsistent}$ = 8.72, $M_{Experiment3*Consistent}$ = 8.42 s,

$M_{Experiment3*Inconsistent}$ = 11.54; $F(1, 45)=0.15$, $P < 0.05$) (see Figure 11). Post hoc Scheffé tests revealed that in Experiment 1, average looking time at the Inconsistent Outcome was not different than average looking time at the Consistent Outcome ($P < 0.59$), while in Experiment 3, average looking time at the Inconsistent Outcome was higher than average looking time at the Consistent Outcome ($P < 0.005$). The average looking time at the Inconsistent Outcome in Experiment 3 was also higher than both the average looking time at the Inconsistent Outcome in Experiment 2 ($P < 0.005$) and the average looking time at the Consistent Outcome in Experiment 2 ($P < 0.001$). All other pairwise comparisons did not detect differences. This pattern of looking times suggests that the Inconsistent Outcome caused an increase in infants looking times only in Experiment 3.

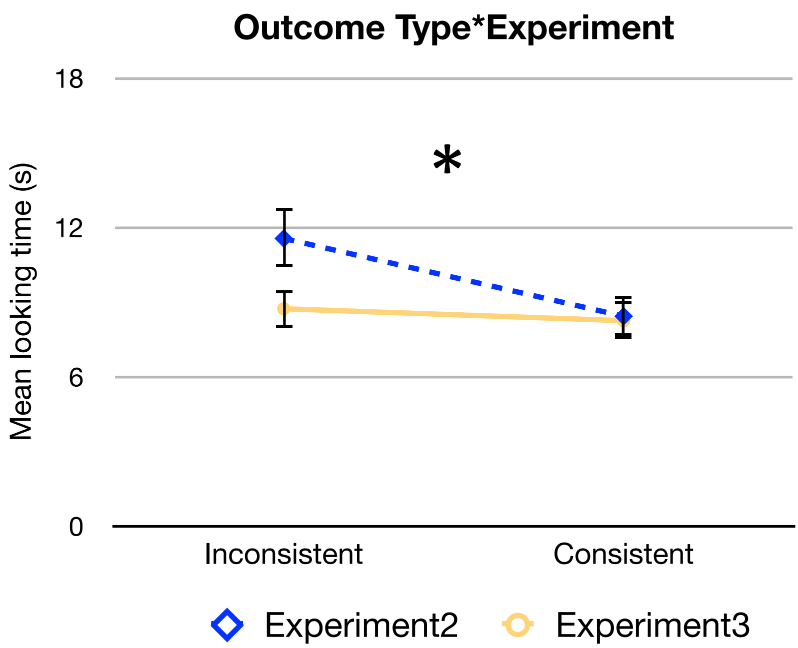

Figure 11. Mean looking times at the Inconsistent Outcome and Consistent Outcome in function of Experiment (Experiment 2 and Experiment 3). Error bars represent SEs.

2.4.3. Discussion

The 12-month-old infants tested in Experiment 3 looked longer at the Inconsistent Outcome than at the Consistent Outcome, a result predicted by the DS hypothesis.

The joint analysis of Experiment 2 and Experiment 3 revealed that 12-month-old infants did react with surprise to the Inconsistent Outcome, although only in Experiment 3. This result shows that 12-month-olds' failure in Experiment 1 did not depend on the lack of ability to perform a disjunctive syllogism. Instead, the results suggest that the procedure used in Experiment 1 did not fairly test infants' ability to detect the Inconsistent Outcome. Experiment 2 involved procedure specific difficulty factors that had prevented the infants from solving the task. Twelve-months-old infants' success in Experiment 3 suggests that they might possess the representational resources required to detect the Inconsistent Outcome of the Scooped Object task.

Unfortunately, it is not possible to establish which difficulty factor mitigated in Experiment 3 (occlusion length, evidence of the ambiguity of the containment event, or both) was the determinant for the failure in Experiment 1. The reason is that the two factors have been manipulated together. However, the results show that the joint mitigation of those two factors was sufficient to allow 12-month-olds to detect the Inconsistent Outcome.

Finally, the detected interaction between the Outcome Type and the objects used revealed that infants reacted with higher looking time at the Inconsistent Outcome when it was presented with the Umbrella/Face pair, but not when presented with the Dinosaur/Flower pair. Such interaction suggests that the employment of the Umbrella/Face pair helped the infants to solve the Scooped Object task. It is possible that the features of that pair, such as, for example, the category contrast HUMAN/NON-HUMAN, supported the encoding and processing of the objects representation across the complex occlusion events that occurred in the task (L.L. Bonatti et al., 2002; L. L. Bonatti et al., 2005).

Taken together, the results of Experiments 1 to 3 suggest that when presented with the Scooped Object task, 19-month-olds react with surprise to the Inconsistent Outcome and, under certain conditions, 12-month-olds do the same.

This pattern of results makes a prima facie case that at least from the age of 12 months, infants might already be able to perform a disjunctive syllogism.

In Experiment 4 (the next experiment) we assessed the robustness of the main effect of the Outcome Type in Experiment 3 by trying to replicate it with a slightly different procedure.

In Experiment 3 infants succeeded in the task only with the Face/Umbrella pair, suggesting that the features of that pair might have helped infants to solve the task. Following this tentative explanation we replaced the Dinosaur/Flower pair used in the previous experiments with a second (new) pair designed to cue the category contrast HUMAN/NON-HUMAN.

The DS hypothesis predicts longer looking time at the Inconsistent Outcome than to the Consistent one.

A further innovation in Experiment 4 is the use of an eye-tracker to measure the looking time at the outcomes with an automatic procedure.

2.5. Experiment 4

2.5.1. Materials and methods

Materials and Methods of Experiment 4 were identical to the ones of Experiment 3, except for the changes indicated in the following sections.

2.5.1.1. Participants

Twenty-three healthy full-term 12-month-old infants were retained for the analyses (11 girls, 12 boys, mean age: 12;09, range: 12;00-12;30). An additional 24 infants were tested but not retained due to fussiness (6), caretakers' interaction (1), equipment failure (3), experimenter error (2), calibration failure (11) or insufficient valid samples (2) (see section 2.2.2.). The participants' parents were contacted by telephone and were given a small present and a certificate of attendance.

2.5.1.2. Materials

2.5.1.2.1. Familiarization movies

In the current experiment we used six familiarization movies (see APPENDIX1 for pictures and detailed descriptions of the content of the familiarization movies).

The familiarization movies F1c and F2c were designed to introduce the functions of the cup and of the occluder when the two "act" separately (i.e. the cup scoops the object only after the removal of the occluder). In each movie, an object (a red puppet in F1c and a red star in F2c) is covered by an occluder and then retrieved once the occluder is removed. Afterwards the cup takes the object, carries it to a different location and releases it (see APPENDIX1, section 7).

The familiarization movies F3c and F4c were designed to introduce the interaction of the functions of the cup and the occluder (i.e. the cup scooped one object from behind the occluder and brought it out). In each movie, an object (the red puppet in F3c and the red star in F4b) is covered by an occluder. Afterwards the cup scoops the object behind the occluder and carries it to a different location. Once the occluder and the cup are removed, the object is released by the cup (see APPENDIX1, section 8).

The familiarization movies F5c and F6c were designed to introduce the interaction of the functions of the cup and the occluder when two objects are hidden behind the occluder, without ever showing the outcome (which object is hidden in the cup). In each movie, two objects (the red puppet and the red star) are covered by an occluder. Afterwards the cup scoops an object behind the occluder and carries it to a different location. Eventually the occluder is removed and an object behind it is found (the red puppet in F7b and the red triangle in F8b), while the other object remains hidden inside the cup (see APPENDIX1, section 9).

In the familiarization movies of the current experiment, participants are never presented with the test objects and the familiarization movies are slightly faster than in Experiment 3.

2.5.1.2.2. Test movies

In the Experiment 4 test movies, the Dinosaur/Flower pair was replaced with a human face and a red flower pair (see Figure 12).

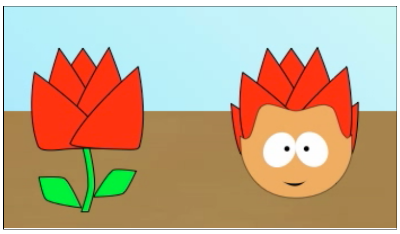

Figure 12. The pair Boy/Flower.

The Boy/Flower pair (we will call the new face "the boy" to distinguish it from the face with the cap in the Face/Umbrella pair) was designed to cue the attribution of the categories HUMAN/ NON-HUMAN. The boy has big eyes and other human-face-like features similar to the face of the Face/Umbrella pair. Furthermore, during its characteristic movement the face moves its gaze and mouth in a human-like manner. The characteristic movement of the flower consists of oscillating left and right several times. The total length of each of the test movies is 17.51 s.

2.5.1.3. Procedure

The procedure of Experiment 4 is identical to the procedure of Experiment 3 except for the use of an eye tracker to record participants' looking behavior, the use of a calibration procedure, the reduction of the Familiarization trials to 6 (one for each movie) and for the way that participants' looking behavior was coded.

In experiment 4 the movies were played on a Tobii T60XL eye tracker (http://www.tobii.com/en/eye-tracking-research/global/products/hardware/tobii-t60xl-eye-tracker/). The window stimuli area (the total area of the movies) is 28.3 x 22.4 cm. The eye tracker was calibrated to each participant before starting the Familiarization Phase. The calibration procedure was carried out by means of the software PsyScope X (http://psy.ck.sissa.it/) and its module TobiiPlus (http://psy.ck.sissa.it/RunTimeInfo/Tobii_and_PsyScope.html) that handled all the interactions with the eye tracker. During the calibration, the participant is sitting on the lap of their caregiver. While colorful images were played on the screen, the distance of the participants from the screen and the height of their eyes were adjusted. The optimal position of the participants' eyes was about 650 mm distance from the screen and at the height of the upper half of the screen. When the participant position is optimal, their attention is drawn to the center of the screen by a central attractor accompanied by cheerful music. When the participant fixates on the central attractor, it disappears and the music stops. Afterwards, a smaller attractor appears for about 2 s on the upper-left extremum point of the screen accompanied by sounds, and samples of pupil and corneal reflection are recorded. The same process is repeated for the upper-right extremum point, for the center of the screen, and for the lower-left and lower-right extrema point. The procedure is repeated until five valid calibration points are obtained. Once the calibration is complete, the participant is presented with a colorful picture accompanied by cheerful music for few seconds. Immediately afterwards, the Familiarization Phase begins.

In the Familiarization Phase of the current experiment, participants saw six movies. We decided to adopt a shorter familiarization than in Experiment 3, since the calibration procedure might increase the overall burden of the experiment. Participants began by viewing two movies that separately introduced the cup and occluder functions. In the third and fourth familiarization trials they saw two movies that introduced the interaction between the cup and the occluder. Eventually, they saw two movies that introduced the test events structure, except for the equivalent of the outcome. See familiarization summary table (for a summary of the familiarization structure see Table 3).

Trial	Movie	Function	Length	Reference
1-2	F1c-F2c, counterbalanced order	familiarization with cup and occluder	2 x 13.48 s	APPENDIX1 section 7
3-4	F3c-F4c, random order	familiarization with interaction of cup and occluder	2 x 11.73 s	APPENDIX1 section 8
5-6	F5c-F6c, counterbalanced order	familiarization with test trial structure	2 x 18.98 s	APPENDIX1 section 9

Table 3. Summary of the structure, functions and length of the familiarization of Experiment 4.

The structure and the procedure of the Test Phase of the current experiment was identical to that of Experiment 4 except for the following adjustments to the administration of the movies:

1. Half of the participants were presented with Consistent or Inconsistent Outcomes in the sequence CICI, while the other half were presented with the sequence ICIC ("I" stands for "Inconsistent Outcome" and "C" stands for "Consistent Outcome").
2. Similarly, the Distance condition was administered either in the sequence CDCD or in the sequence DCDC ("C" stands for "Close" and "D" stands for "Distant").
3. Each participant was tested two times with each object pair. The order of presentation of the pair was either $C_1C_2C_1C_2$ or $C_2C_1C_2C_1$ ("C_1" stands for one pair, "C_2" stands for the other pair).
4. Each participant was presented with all four levels of the variable Object (the type of object retrieved inside the cup). Across participants, the trial number was crossed with which object was found inside the cup.

In Experiment 4, the off-line looking time was not measured by manually coding recorded videos of a participant performance. Instead, in the current experiment the eye-tracker data were used to automatically calculate the looking time. The eye tracker sampled the participant's gaze at 60 Hz. Thus, the looking time (in seconds) in a trial was calculated as one sixtieth of the amount of gaze points at the movie frame that occurred in a trial. As in the previous experiments, the exact moment the trial ended was coded as occurring when either

2 s of looking outside of the movie frame were recorded or a total of 30 s of looking within the movie frame were recorded.

2.5.2. Results

Rejection criteria were identical to those of previous experiments. The trial rejection rate due to the filter based on the deviation from the median was 4.5% of the Inconsistent Outcome trials and 11.1% of the Consistent Outcome trials.

Preliminary analyses (see APPENDIX2 for the complete results of the analyses) detected no interaction effects of Outcome Type with Pair (Star/Triangle vs Dinosaur/Flower vs Face/Umbrella) or Distance (Close vs Distant). Therefore the data were collapsed across Pair and Distance in the subsequent analysis.

For each participant, two mean looking time scores were calculated by averaging the looking time trial scores by Outcome Type (Consistent or Inconsistent). To test for a difference in performance between the two outcome types an ANOVA with Outcome Type as a within-participants factor and Participant as random variable was performed on mean looking time scores. The analyses detected an effect of the Outcome Type ($M_{Consistent}$ = 6.3 s, $M_{Inconsistent}$ = 7.6 s; $F(1, 24) = 4.49$, $P < 0.05$), showing that the 12-month-olds looked at the Inconsistent Outcome longer than the Consistent Outcome (see Figure 13). A Wilcoxon Signed-ranks test confirmed the main effect of the Outcome Type ($z = -1.82$, $P = 0.07$, two-tailed).

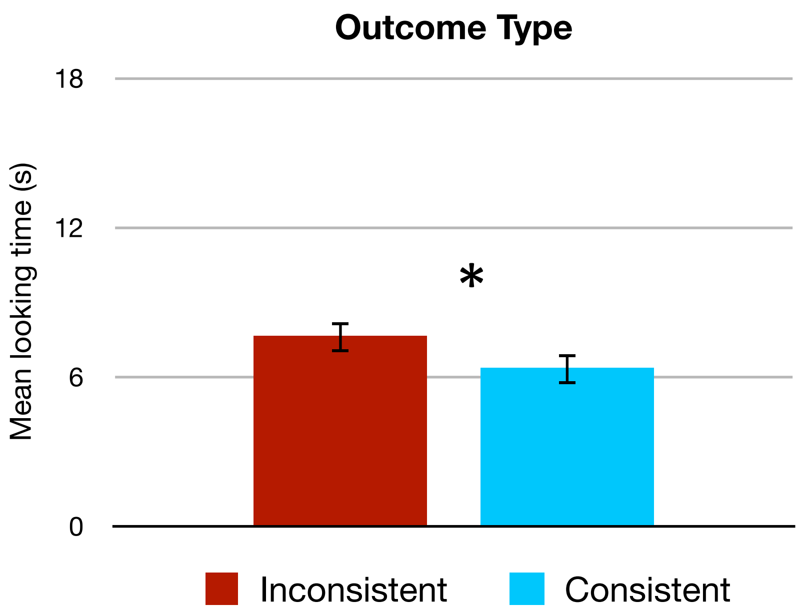

Figure 13. Twelve-month-olds' mean looking times at the Inconsistent Outcome and Consistent Outcome in Experiment 4. Error bars represent SEs.

In order to compare the effect of Outcome Type on participants' looking times in Experiment 2 and Experiment 3, the mean looking times per condition of the participants of the two experiments were entered in a two-way repeated measures ANOVA with Outcome Type as a within-participants factor, Experiment as a between-participants factor and Participant as a random factor nested in Experiment. There was a main effect of Outcome Type ($M_{Possible}$ = 7.3 s, $M_{Consistent}$ = 9.6 s; $F(1, 44) = 12.53$, $P = 0.001$) and a main effect of Experiment ($M_{Experiment3}$ = 10 s, $M_{Experiment4}$ = 6.9 s; $F(1, 44) = 9.39$, $P < 0.005$). No interaction between Outcome Type and Experiment was detected ($M_{Experiment3*Consistent}$ = 8.42 s, $M_{Experiment3*Inconsistent}$ = 11.54 s, $M_{Experiment4*Consistent}$ = 6.3 s, $M_{Experiment4*Inconsistent}$ = 7.56 s; $F(1, 44) = 2.23$, $P < 0.142$); (see Figure 14). This pattern of looking times suggests that the Inconsistent Outcome caused an increase in infants' looking times in both Experiments 3 and 4.

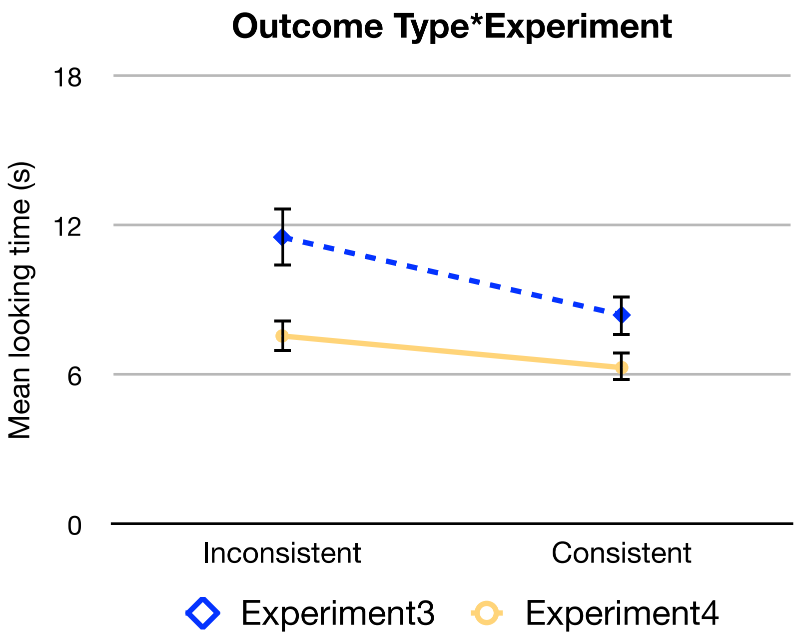

Figure 14. Mean looking times at the Inconsistent Outcome and Consistent Outcome in function of Experiment (Experiment 3 and Experiment 4). Error bars represent SEs.

2.5.3. Discussion

The 12-month-old infants tested in Experiment 3 looked at the Inconsistent Outcome for longer than at the Consistent Outcome, a result predicted by the DS hypothesis.

Thus, Experiment 4 replicates the result of Experiment 3 and shows its robustness.

No interaction between the Outcome Type and the pair used was detected. This result shows that in Experiment 4, infants' higher looking time at the Inconsistent Outcome (compared with the

Consistent one) was not strongly modulated by the pair used to produce the outcome. The absence of an interaction of Pair and Outcome Type suggests that in Experiment 4 infants succeed in the task with both the new Boy/Flower pair and in the Face/Umbrella pair.

The joint analysis of Experiments 3 and 4 detected longer looking time at the Inconsistent Outcome than at the Consistent Outcome and no interaction between the Outcome Type and the experiments, showing that in both experiments infants were able to detect the Inconsistent Outcome. Furthermore, infants looking time was generally higher in Experiment 3 than in Experiment 4. One possible explanation of this difference is a product of the procedure we adopted starting from Experiment 4. The shorter average looking time in Experiment 4 might be due to the joint effect of the eye tracker precision in detecting gaze points that occurred outside of the stimuli window and the tendency of the eye tracker to lose a small percentage of gaze points. On the one hand, the possible higher precision of the eye tracker in detecting whether a gaze point was outside the stimuli window might have fostered the recognition of when a participant had looked away from the stimuli window for at least 2 s, and thus might have resulted in a shortened looking time. On the other hand, the percentage of gaze points lost might be reflected by the lower average looking time, as a result of the procedure we adopted to measure looking time in Experiment 4. Indeed, the looking time in a trial was calculated as a linear function of the number of gaze points directed toward the stimuli window recorder in the Test Phase of the trial.

2.6. General discussion of Experiments 1 to 4

Experiments 1 to 4 suggest that when presented with the Scooped Object task, 19-month-olds reacted with surprise to the Inconsistent Outcome and, under certain conditions, 12-month-olds did the same. Infants' and toddlers' responses to the Scooped Object task are explained by the DS hypothesis: infants and toddlers might have represented disjunctively the epistemic alternatives compatible with the ambiguous containment and inferred by disjunction elimination which object was inside the cup.

However, the conclusion that infants or toddlers are able to perform a disjunctive syllogism may be premature. An alternative explanation proposes that participants detected an incongruence in the Inconsistent Outcome not because they performed a logical inference, but because they tried to identify the object inside the cup only after it is revealed.

For example, it is possible that when participants initially saw the two objects they formed a memory model of them (a representation of the objects in the scene) and maintained it in memory until the objects were retrieved (Carey & Xu, 2001; Fei Xu, 2007; J. Zosh & Feigenson, 2009). The memory model might have consisted of one object representation for each one of the attended objects, Object A and Object B. Each object representation encoded the visual properties, category membership of its denotation (i.e. the object it refers to). Each time an object was retrieved (i.e. it became fully visible again), the object is identified with one of the object representations of the memory model, based on its visual properties or category membership.

While in the Consistent Outcome two objects with different looks were retrieved (e.g. a boy and an umbrella), in the Inconsistent Outcome two objects with the same look were retrieved (e.g. a boy and a boy).

As a consequence, participants might have been surprised by the Inconsistent Outcome, for example because of the appearance of a second object of one category (e.g. a second boy).

This kind of object identification strategy does not require logical representations. No disjunctive representation and no disjunction elimination is required to apply it. The strategy instead requires representational skills that have been documented and studied infants in the second year of life: the ability to store in memory representations of attended objects that encode their features and spatial location and the ability to use those representations to individuate newly appeared objects (Carey & Xu, 2001; F Xu & Carey, 1996; Fei Xu, 2007; J. Zosh & Feigenson, 2009). That is, the ability to use featural or spatiotemporal information to decide whether a newly perceived object is the same or a different individual than an object whose representation is stored in memory.

Thus, an object mapping strategy like the one sketched above might be an alternative explanation of the results of Experiments 1 to 4.

In Experiments 5 and 6 we tested toddlers and infants with a new task designed to establish whether, in dealing with the ambiguous containment, they rely on a disjunctive syllogism or else on this kind of object identification process.

3. EXPERIMENTS 5 AND 6: THE HIDDEN OBJECT TASK

3.1. Introduction

The Hidden Object task is a new version of the Scooped Object task. The Hidden Object task is specifically designed to test whether, in dealing with the ambiguity of the scooping event, participants perform a disjunctive syllogism or if they wait for the object inside the cup to reappear to identify it.

In the Hidden Object task infants are presented with movies that are identical to the movies of the Scooped Object task up to the moment when the cup lands in the right region of the screen with an object inside, after having visited the region hidden by the occluder. In the Hidden Object task the scooping event is produced in exactly the same way as in the Scooped Object task: two objects are hidden behind an occluder and one of the two is scooped out by a cup that reveals enough of the object to know that an object is inside it, but not which object.

The first change in the task is that the occluder is never removed. Instead one of the two objects, e.g. Object A, exits form the right edge of the occluder, stays entirely visible to the viewer for about 1 s and afterwards returns behind the occluder. (see Figure 15, blue frame). The second change is the conclusion of the movies. In the Hidden Object task the object hidden inside the cup is never revealed. Instead for a second time an object exits from the right edge of the occluder. In the Consistent Outcome an object that looks identical to the object that has just returned behind the occluder, e.g. identical to Object A, emerges from the occluder (see Figure 15, green frame). In the Inconsistent Outcome an object emerges from the occluder that looks identical to the other object, e.g. identical to Object B (see Figure 15, red frame).

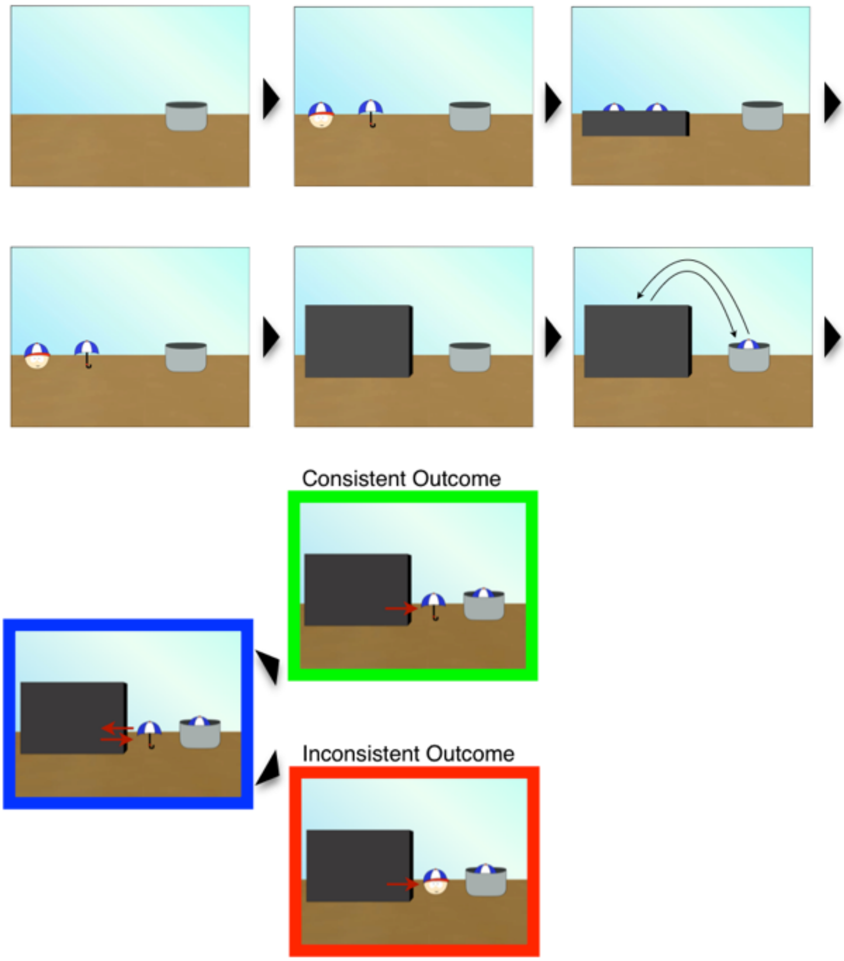

Figure 15. Schema of the Hidden Object task. The occluder is never removed, instead one object exits form the right edge of the occluder and afterwards returns behind it. (blue frame). The object hidden inside the cup is never revealed, instead for a second time an object exits from the right edge of the occluder. In the Consistent Outcome it looks identical to the object that has just returned behind the occluder (green frame), while in the Inconsistent Outcome it looks identical to the other object (red frame).

The DS hypothesis again predicts longer looking time at the Inconsistent Outcome than at the Consistent Outcome. Somebody equipped with the ability to perform a disjunctive syllogism can infer what object has to be inside the cup by witnessing the first object that exits from the side of the occluder. That is, someone may reason that since *either* Object A was inside the cup *or* Object B was inside the cup, and Object A is behind the occluder (and thus *not* inside the cup), then the object inside the cup has to be Object B. Crucially, the rational expectation that Object B is inside the cup (and thus Object A is behind the occluder) is incompatible with the Inconsistent Outcome. In the Inconsistent Outcome an object that looks identical to the one that by logic has to be inside the cup (e.g. identical to Object B) exits from behind the occluder. In contrast, the rational expectation that Object B is inside the cup is compatible with the Consistent Outcome. In the Consistent Outcome an object that looks identical to the object has just has just returned behind the occluder (e.g. identical to Object A) exits from behind the occluder.

In contrast, a strategy based on identifying the objects based on their appearance once they have reappeared will not detect a mismatch in the Inconsistent Outcome of the Hidden Object task.

In the Possible Outcome an object looking like Object A exits and returns behind the occluder, and then an object looking like Object A exits from the Occluder again. Each time the appeared object will be plausibly identified with Object A, because of its appearance (Leslie, Xu, Tremoulet, & Scholl, 1998; F Xu & Carey, 1996). In the Impossible Outcome the first appearing object looks like Object A while the second looks like Object B, thus it is plausible that the two objects will be identified respectively with Object A and Object B, because of their appearance. No mismatch in the

Inconsistent Outcome should plausibly be detected by identifying the visible objects based on their appearance, since in the outcome two objects with the same appearance do not appear.

In summary, while the DS hypothesis predicts longer looking times at the Inconsistent Outcome than at the Consistent Outcome in the Hidden Object task, the strategy based on identifying objects when they emerge from occlusion does not. In Experiment 5 we test these predictions with a group of 19-month-old toddlers.

3.2. Experiment 5

3.2.1. Materials and methods

3.2.1.1. Participants

Twenty-four healthy full-term 19-month-old toddlers were retained for the analyses (12 girls, 12 boys, mean age: 19;08, range: 18;15-20;02). An additional 12 toddlers were tested but not retained due to fussiness (4), caretakers' interaction (1), experimenter error (1) or refusal to begin the experiment (6). The participants' parents were contacted by telephone and were given a small present and a certificate of attendance.

3.2.1.2. Materials

The materials of Experiment 5 were identical to those of Experiment 4 except for the changes indicated in the following sections.

3.2.1.2.1. Familiarization movies

In Experiment 5 we used six brand new familiarization movies and one movie from Experiment 1 (see APPENDIX1 for pictures and detailed descriptions of the content of the familiarization movies).

The familiarization movies F1d and F2d were designed to introduce the functions of the cup and of the occluder when the two "act" separately (i.e. the cup scoops the object only after the removal of the occluder). In each movie, an object (a blue puppet in F1b and a blue flower in F2b) is taken by the cup, carried to a different location and eventually released. Afterwards, the occluder is raised over the left empty region of the screen and is then removed (see APPENDIX1, section 10).

The familiarization movies F3d and F4d were designed to introduce the interaction of the functions of the cup and the occluder (i.e. the cup scooped one object from behind the occluder and brought it out). In each movie, an object (the blue puppet in F3b, the blue flower in F4b) is covered by an occluder. Afterwards, the cup scoops the object behind the occluder and carries it to a different location. The occluder is not removed and when the cup is removed, the object is released by the cup (see APPENDIX1, section 11).

The familiarization movies F5d and F6d were designed to show that either one of the two objects can exit from the right edge of the occluder, irrespective of the position that it was initially occupying when the occluder covers the two objects. In each movie, two objects (the blue puppet and the blue flower) are covered by an occluder. Afterwards, one of the two objects (the blue puppet in F5d and the blue flower in F6d) exits from the right edge of the occluder (see APPENDIX1, section 12).

The movie F4a (already used in Experiment 1) was designed to introduce all test objects. The six objects lie side by side and move one after the other to attract the viewer's attention (see APPENDIX1, section 13).

3.2.1.2.2. Test movies

In the Test Movies of Experiment 5 we used the same test pairs used with Experiment 1. That is, in the Experiment 5 test movies we used the Star/Triangle pair, the Dinosaur/Flower pair and the Face/ Umbrella pair. Since the 19-month-olds tested in Experiment 1 were able to solve the Scooped Object task with these pairs, we expect toddlers to succeed in the Hidden Object task with the same object pairs. To preserve a counterbalanced design, we presented each participant with six test trials, as in Experiment 1.

The Test Movies of Experiment 5 were identical to the Test Movies of Experiment 4 up to the moment when the cup lands in the right region of the screen with an object inside, after having visited the region hidden by the occluder. Afterwards, one of the two objects exits from the right edge of the occluder and reaches the middle point between the occluder and the cup, in 0.25 s. From this position the object executes its characteristic movement accompanied by a sound. The star and the triangle pulsate repeatedly, the flower and the umbrella oscillate left and right several times, the face shows a smile, the dinosaur executes a jump and tilts its head up and down. Two s after the moment when the object has reached the middle point between the occluder and the cup, the object returns behind the occluder, a movement that lasts about 0.25s. About 1 s after the moment when the object has returned behind the occluder, an object exits again from the side of the occluder and moves to the middle point between the occluder and the cup. In half of the test movies the second object that exits from behind the occluder looks identical to the first object that had exited from behind the occluder, this is the Consistent Outcome. In the other half of the test movies, the second object that exits from behind the occluder looks identical to the other object of the pair, this is the Inconsistent Outcome (see Figure 16). The total length of each test movie is 17.93 s. The time elapsed from the moment the two objects are completely covered by the occluder to the moment that an object exits from behind the occluder for a second time is about 8 s.

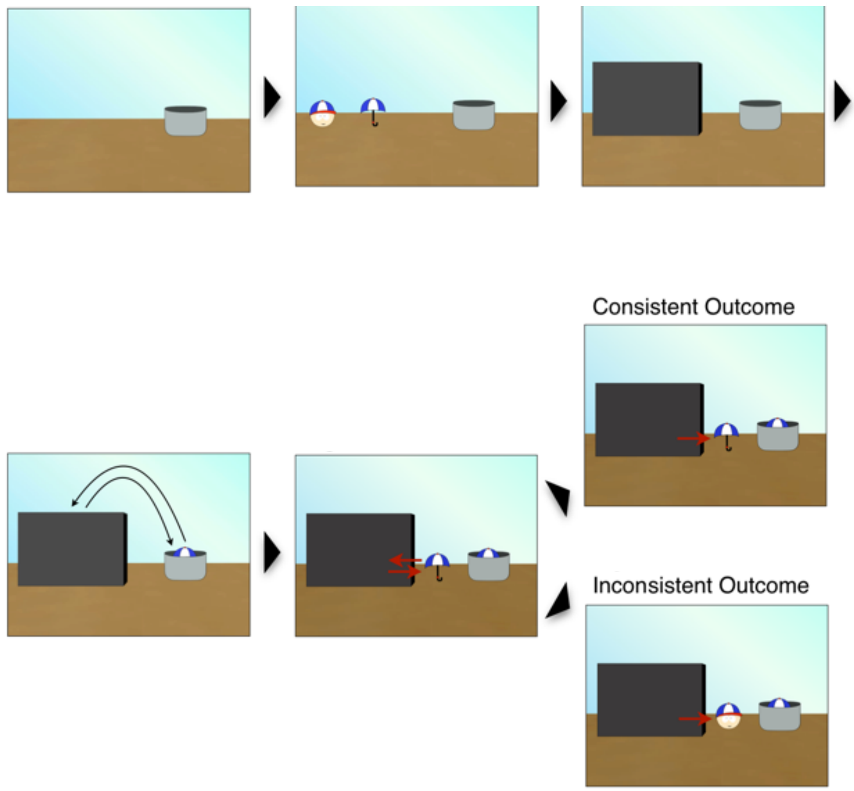

Figure 16. The two types of outcomes in the Hidden Object task. In the Consistent Outcome the second object that exits from behind the occluder looks identical to the first object that had exited from behind the occluder. In the Inconsistent Outcome the second object that exits from behind the occluder looks identical to the other object of the pair.

Across the 24 test movies, the following variables were crossed:

1. Outcome Type. Two levels: Consistent or Inconsistent.
2. Pair. Three levels: Star/Triangle or Flower/Dinosaur or Umbrella/Face.
3. Object; the type of the second object that exits from behind the occluder. Six levels: Star, Triangle, Flower, Dino, Umbrella or Face; nested in Pair.

4. Distance; the distance from the cup at the initial moment of the movie of the second object that exits from behind the occluder. Two levels: Distant or Close.

Within each pair, the position of the objects at the beginning of the movie was inverted in half of the movies.

3.2.1.3. Procedure

In the Familiarization Phase of the current experiment participants saw seven movies. Participants began by viewing two movies that separately introduced the cup and the occluder functions. In the third and fourth familiarization trials they saw two movies that introduced the interaction between the cup and the occluder. In the fifth and sixth familiarization trials they saw two movies that show that either object can exit from the occluder (one per movie). Eventually, they saw a movie that introduced all test objects (for a summary of the familiarization structure see Table 4).

Trial	Movie	Function	Length	Reference
1-2	F1d-F2d, counterbalanced order	familiarization with cup and occluder	2 x 13.48 s	APPENDIX1 section 10
3-4	F3d-F4d, counterbalanced order	familiarization with interaction of cup and occluder	2 x 15.36 s	APPENDIX1 section 11
5-6	F5d-F6d, counterbalanced order	familiarization with the exit of objects from the occlude	2 x 14.79 s	APPENDIX1 section 12
7	F4a	familiarization with test objects	1 x 13 s	APPENDIX1 section 3

Table 4. Summary of the structure, functions and length of the familiarization of Experiment 5.

The familiarization movies of the current experiment were designed to be similar to those of Experiment 1 in that the movies never show the cup event when two objects are present in the scene and the last movie shows all the test objects together in one scene. Furthermore, we decided to replace the familiarization objects used in Experiment 1 (bear and red car) with a pair of objects that have an identical upper part, like we did in Experiments 3 and 4. The new familiarization objects were a blue puppet and a blue flower. We decided to use a brand new familiarization pair, and not the Red-Puppet/Red-Star pair, since in the current experiment we used the test pair Triangle/Star.

The structure and the procedure of the Test Phase of the current experiment was identical to that of Experiment 4 except for the number of movies presented, which was six, and the following adjustments to the administration of the movies:

1. Half of the participants were presented with Consistent or Inconsistent Outcomes in the sequence CIICCI, while the other half were presented with the sequence ICCIIC ("I" stands for "Inconsistent Outcome" and "C" stands for "Consistent Outcome").
2. Similarly, the Distance condition was administered either in the sequence CDDCCD or in the sequence DCCDDC ("C" stands for "Close" and "D" stands for "Distant").
3. Each participant was tested two times with each object pair. The order of presentation of the pair was either $C_1C_2C_3C_1C_2C_3$ or $C_2C_3C_1C_2C_3C_1$ or $C_3C_1C_2C_3C_1C_2$ ("C_1" stands for one pair, "C_2" stands for the second pair, "C_3" stands for the remaining pair).
4. Each participant was presented with all four levels of the variable object (the type of object retrieved inside the cup). Across participants, the trial number was crossed with which object was found inside the cup.

3.2.2. Results

Rejection criteria were identical to those of the previous experiments. The trial rejection rates due to the filter based on the deviation from the median was 10.2% of the Inconsistent Outcome trials and 6.9% of the Consistent Outcome trials.

Preliminary analyses (see APPENDIX2 for complete results of the analyses) detected no interaction effects of Outcome Type with Pair (Star/Triangle vs Dinosaur/Flower vs Face/Umbrella) or Distance (Close vs Distant). Therefore the data were collapsed across Pair and Distance in the subsequent analysis.

For each participant, two mean looking time scores were calculated by averaging the looking time trial scores by Outcome Type (Consistent or Inconsistent). To test for a difference in performance between the two outcome types an ANOVA with Outcome Type as a within-participants factor and Participant as random variable was performed on mean looking time scores. The analyses detected an effect of the Outcome Type ($M_{Consistent}$ = 5 s, $M_{Inconsistent}$ = 6.3 s; $F(1, 23) = 8.2$, $P < 0.01$), showing that the 19-month-olds looked at the Inconsistent Outcome for longer than the Consistent Outcome (see Figure 17).

A Wilcoxon Signed-ranks test confirmed the main effect of the Outcome Type ($z = -2.57$, $P < 0.05$, two-tailed).

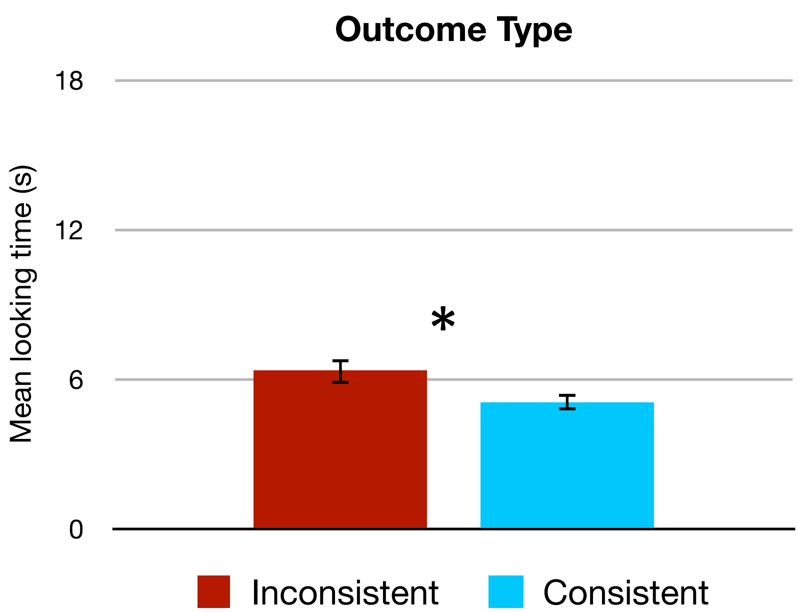

Figure 17. Nineteen-month-olds' mean looking times at the Inconsistent Outcome and Consistent Outcome in Experiment 5. Error bars represent SEs.

3.2.3. Discussion

The 19-month-old toddlers tested in Experiment 5 looked at the Inconsistent Outcome for longer than at the Consistent Outcome. That is, after seeing that Object A exited from and returned behind the occluder, toddlers looked for longer at the outcome where one object that looked identical to Object B exited from behind the occluder, than at the outcome where an object that looked identical to Object A exited from behind the occluder.

This reaction is not predicted by the proposal that toddlers simply try to identify the objects that emerge from the occlusion based on their appearance.

These results are instead predicted by the DS hypothesis, according to which, after seeing that Object A was behind the occluder, toddlers inferred by disjunction elimination that Object B was inside the cup (and thus expected that only Object A was behind the occluder).

In Experiment 6 we presented 12-month-old infants with the Hidden Object task. If, as proposed by the DS hypothesis, infants infer that Object B is inside the cup by witnessing that Object A is outside the cup, then they will look at the Inconsistent Outcome for longer than at the Consistent Outcome. On the other hand, if, as proposed by the object individuation explanation discussed above, infants do not infer that Object B is inside the cup once they have seen that Object A is outside the cup, but wait until the objects reappear to identify them, then they will not look at the Inconsistent Outcome for longer than at the Consistent Outcome.

3.3. Experiment 6

3.3.1. Materials and methods

The materials and methods of Experiment 6 were identical to those of Experiment 5 except for the changes indicated in the following sections.

3.3.1.1. Participants

Twenty-four healthy full-term 12-month-old infants were retained for the analyses (13 girls, 11 boys, mean age: 12;02, range: 11;22-12;13). An additional 28 toddlers were tested but not retained due either to fussiness (10), caretakers' interaction (5), experimenter error (4), calibration failure (1), refusal to begin the experiment (4). or insufficient valid samples (4) (see section 2.2.2.). The participants' parents were contacted by telephone and were given a small present and a certificate of attendance.

3.3.1.2. Materials

3.3.1.2.1. Familiarization movies

In the current experiment we used six brand new familiarization movies and two movies from Experiment 3 (see APPENDIX1 for pictures and detailed descriptions of the content of the familiarization movies).

The familiarization movies F1b and F2b were designed to introduce the functions of the cup and of the occluder when the two "act" separately (i.e. the cup scoops the object only after the removal of the occluder). In each movie, an object (a red puppet in F1b and a red star in F2b) is covered by an occluder and then retrieved once the occluder is removed. Afterwards, the cup takes the object, carries it to a different location and eventually releases it (see APPENDIX1, section 14).

The familiarization movies F3e, F4e, F5e and F6e were designed to introduce the interaction of the functions of the cup and the occluder (i.e. the cup scooped one object from behind the occluder and brought it out). It also presents to the infants the way each test object looks when partially hidden inside the cup. In each movie, an object is covered by the occluder (the worm in F3e, the ball in F4e, the umbrella in F5e and the face in F6e; see section 3.3.1.2.2. for a description of the objects "Worm" and "Ball"). The cup then scoops the object behind the occluder and carries it to a different location. The occluder is not removed and, when the cup is removed, the object is released by the cup (see APPENDIX1, section 15).

The familiarization movies F7e and F8e were designed to introduce the interaction of the functions of the cup and the occluder when two objects are hidden behind the occluder, without ever showing the outcome (which object is hidden in the cup). The presentation of the scooping event in the familiarization phase may support the recognition of the ambiguity of the containment in the test. In each movie, two objects (the red puppet and the red star) are covered by an occluder. The cup then scoops an object behind the occluder and carries it to a different location. Eventually the occluder is removed and an object behind it is found (the red puppet in F7b and the red triangle in F8b), while the other object remains hidden inside the cup (see APPENDIX1, section 16).

3.3.1.2.2. Test movies

In the Test Movies of Experiment 6 we used the Face/Umbrella pair and a new pair comprising a red worm with eyes and a multicolored ball, the Worm/Ball pair (see Figure 18).

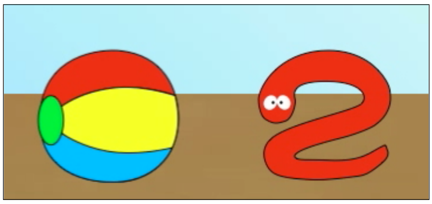

Figure 18. The pair Worm/Ball.

The Worm/Ball pair were designed to cue the attribution of the categories contrast AGENT/NON-AGENT, a categories contrast that has been shown to help object individuation (Surian & Caldi, 2010). When the ball initially enters the scene by falling from the upper edge of the screen frame, it bounces several times while rotating on its center. Its characteristic movement consists of rotating on its center. Whenever it moves horizontally to exit from or return behind the occluder, it rotates in the same direction of the movement. The worm exhibits cues of agency (Surian & Caldi, 2010): it has two big eyes and executes complex articulatory, non-rigid movements both as its characteristic movement and whenever it moves horizontally.

The Test Movies of Experiment 6 have the same structure as the Test Movies of Experiment 5. The total length of each test movie is 19.43 s. The time elapsed from the moment the two objects are completely covered by the occluder up to the moment that an object exits for a second time from behind the occluder is about 9 s.

3.3.1.3. Procedure

In the Familiarization Phase of the current experiment participants saw seven movies. Participants began by viewing two movies that separately introduced the cup and occluder functions. From the third to the sixth familiarization trials they saw four movies that introduced the interaction between the cup and the occluder. Finally, they saw two movies that introduced the test events structure, except for the equivalent of the outcome (for a summary of the familiarization structure see Table 5).

Trial	Movie	Function	Length	Reference
1-2	F1d-F2d counterbalanced order	familiarization with cup and occluder	2 x 13.48 s	APPENDIX1 section 13
3-6	F3d-F6d random order	familiarization with interaction of cup and occluder and with test objects.	4 x 13.86 s	APPENDIX1 section 14
7-8	F7d-F8d counterbalanced order	familiarization with test trial structure	2 x 17.98 s	APPENDIX1 section 15

Table 5. Summary of the structure, functions and length of the familiarization of Experiment 6.

The Familiarization movies of the current experiment were designed to resemble as much as possible the Familiarization movies of Experiment 3. We choose to do so as the familiarization procedure of Experiment 3 is likely to have played a crucial role in infants' success in the Scooped Object task (see sections 2.3.3. and 2.3.4) The structures of movies F3e to F8e match those of the corresponding movies of Experiment 3, except for two aspects. First, the occluder is never removed, since the movies are designed to familiarize the

Hidden Object task. Secondly, which object is not inside the cup is shown by having it exit from the occluder (in F7e and F8e).

The structure and the procedure of the Test Phase of the current experiment was identical to that of Experiment 5 except for the number of movies presented, which was four, and the following adjustments to the administration of the movies:

1. Half of the participants were presented with Consistent or Inconsistent Outcomes in the sequence CIIC, while the other half were presented with the sequence ICCI ("I" stands for "Inconsistent Outcome" and "C" stands for "Consistent Outcome").
2. Similarly, the Distance condition was administered either in the sequence CDDC or in the sequence DCCD ("C" stands for "Close" and "D" stands for "Distant").
3. Each participant was tested two times with each object pair. The presentation order of the pair was either $C_1C_2C_1C_2$ or $C_2C_1C_2C_1$ ("C_1" stands for one pair, "C_2" stands for the other pair).
4. Each participant was presented with all six levels of the variable Object (the type of object retrieved inside the cup). The order of presentation of the object type was counterbalanced.

The trial rejection rate due to the filter based on the deviation from the median was 9.5% of the Inconsistent Outcome trials and 20.4% of the Consistent Outcome trials.

3.3.2. Results

Preliminary analyses (see APPENDIX2) detected no interaction effects of Outcome Type with Pair (Star/Triangle vs Dinosaur/Flower vs Face/Umbrella) or Distance (Close vs Distant). Therefore the data were collapsed across Pair and Distance and in the subsequent analysis.

For each participant two mean looking time scores were calculated by averaging looking time trial scores by Outcome Type (Consistent or Inconsistent). To test for a difference in performance between the two outcome types, an ANOVA with Outcome Type as a within-participants factor and Participant as random variable was performed on mean looking time scores. The analyses detected an effect of the Outcome Type ($M_{Consistent}$ = 4.2 s, $M_{Inconsistent}$ = 6.1 s; $F(1,23)=11.05$, $P<0.005$), showing that the 12-month-olds looked at the Inconsistent Outcome for longer than the Consistent Outcome. (see Figure 19). A Wilcoxon Signed-ranks test confirmed the main effect of the Outcome Type (z = -2.86, P < 0.005, two-tailed)

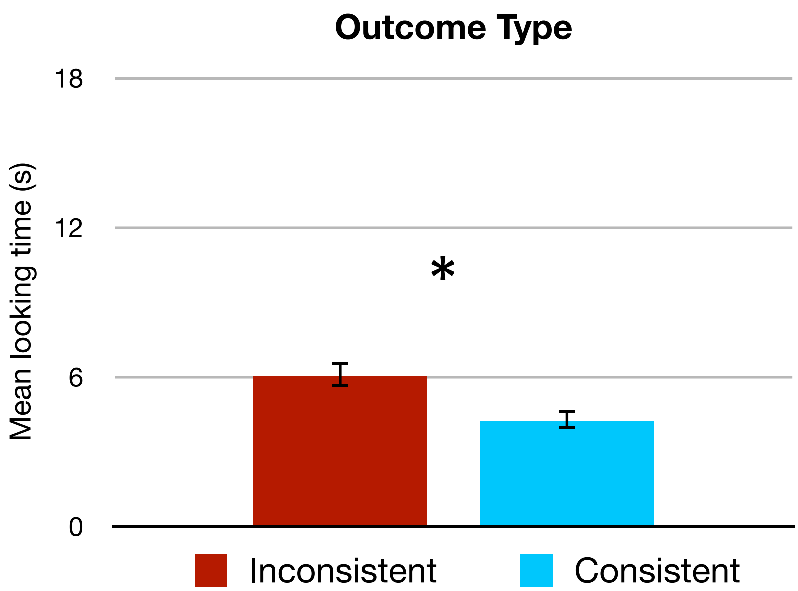

Figure 19. Twelve-month-olds' mean looking times at the Inconsistent Outcome and Consistent Outcome in Experiment 6. Error bars represent SEs.

3.3.3. Discussion

The 12-month-old infants tested in Experiment 5 looked at the Inconsistent Outcome for longer than at the Consistent Outcome.

This reaction is not predicted by the proposal that infants simply try to identify the objects based on their appearance once they have emerged from the occlusion.

The result is instead predicted by the DS hypothesis, according to which, after seeing that Object A was behind the occluder, toddlers inferred by disjunction elimination that Object B was inside the cup.

The absence of interaction between the pair used and the Outcome Type suggests that infants were able to detect the Inconsistent Outcome with both of the pairs.

3.4. General discussion of Experiments 5 and 6

Both toddlers' and infants' looking time responses in the Hidden Object task suggest that they were surprised by the Inconsistent Outcome. Surprise in response to the exit of an object looking like Object B from the occluder, after having seen that an object looking like Object A is behind the occluder (i.e. in response to the Inconsistent Outcome) suggests that participants might have inferred that, since Object A is behind the occluder, Object B has to be inside the cup, as proposed by the DS hypothesis.

In contrast, the proposal that toddlers simply tried to identify the objects that emerged from the occlusion based on their appearance does not predict surprise to the Inconsistent Outcome. In fact, in the Inconsistent Outcome of the Hidden Object task two objects of the same type never emerged.

In the next sections we will discuss three alternative explanations for the results of Experiments 5 and 6.

3.4.1 Novelty effect

Participants' higher looking times at the Inconsistent Outcome of the Hidden Object task may have originated from elements of novelty that are present in the Inconsistent Outcome and absent in the Consistent Outcome (Munakata, 2000). That is, the higher looking time in the Incongruent Outcome condition is not caused by the detection of an inconsistence in the Inconsistent Outcome, but instead by the higher novelty of that outcome compared to the Congruent Outcome.

In fact, in the Hidden Object task, the Inconsistent Outcome does contain elements of novelty that are absent in the Possible

Outcome. In the Consistent Outcome condition the participants witnessed one object exit from and return behind the occluder and afterwards an object exit from behind the occluder, with the two objects being identical. In contrast, in the Inconsistent Outcome condition the two objects that sequentially exited from behind the occluder had very different features. Thus, in the Inconsistent Outcome, with the second appearance of an object, an observer can detect (relatively) new visual properties, a (relatively) new category and a (relatively) new individual (since the features of the second object suggest that it is a distinct individual from the first object). As a consequence, it is possible that participants looked at the Inconsistent Outcome for longer because their attention was attracted by those elements of novelty.

However, it is not clear how an explanation based on novelty effect may account for the results of the Scooped Object task (Experiments 1 to 4). For in the Inconsistent Outcome of the Scooped Object task, the two objects that appear sequentially are identical, while it is in the Congruent Outcome that the two objects have different visual properties and belong to two different categories. Thus, the Congruent Outcome is more novel. In fact, tested with the Scooped Object task, infants and toddlers looked at the outcome with less elements of novelty for longer (the Incongruent Outcome). As a result, a novelty effect alone might be insufficient to explain the looking time results of all six experiments.

3.4.2. Comparison of sets numerosities

The last alternative explanation to be discussed accounts for the results of Experiments 5 and 6 by noticing that the Inconsistent

Outcome of the Hidden Object task, but not the Consistent Outcome, offers the viewer evidence of the presence of a third object in the scene.

A viewer presented with the Hidden Object task who recognizes that an object is hidden inside the cup might be able to count two distinct objects in the Consistent Outcome (the object in the cup and an object looking like Object A that exits, returns and exits again from behind the occluder) and up to three distinct objects in the Inconsistent Outcome (the object in the cup, an object looking like Object A that exits and returns behind the occluder and an object looking like Object B that exits from then occluder).

Infants in their second year of life might have the ability to compare the numerosities of two small sets of objects (Feigenson & Carey, 2003). Therefore, if infants and toddlers expected that the number of objects present in the outcome has to be the same as the number of objects seen at the beginning of the movie, then they might be surprised by the Inconsistent Outcome (while not by the Consistent Outcome) because of the higher numerosity of the final set of objects.

In principle, it seems possible to compare the numerosities of the two sets without inferring exactly which object is inside the cup. Therefore, it seems possible that infants and toddlers have performed such a numerical comparison without performing a disjunctive syllogism.

Thus, a strategy based on this sort of numerical comparison might be a valid alternative to the hypothesis that infants and toddlers performed a disjunctive syllogism in dealing with the Hidden Object task (that is, in Experiments 5 and 6).

However, it is not clear how a numerical comparison of this kind might detect the Inconsistent Outcome of the Scooped Object task (that is, Experiments 1 to 4). For in the Scooped Object task the numerosity of the set of objects initially seen and the numerosity of the set of those in the Inconsistent Outcome is the same.

4. EYE-MOVEMENTS ANALYSES

4. 1. Beyond cumulative looking time

In the discussion of Experiments 1 to 6 we individuated different strategies that might have supported the detection of the Inconsistent Outcome. One strategy is based on the logical inference attributed by the DS hypothesis (see sections 2.1. and 3.1.). The others are the object identification process and the set comparison that might explain the results of Experiments 1 to 4 and Experiments 5 to 6, respectively (see sections 2.6. and 3.4.2.). Those strategies might require very distinct abilities. In particular, the strategy based on the disjunctive syllogism (but not the other two) depends on the possession of logical representations.

Nevertheless, the looking time results of Experiments 1 to 6 may be insufficient to clarify which strategy participants actually employed to solve our task. For, although very different, the strategies make similar predictions in terms of cumulative looking time response to the outcomes.

To find behavioral evidence that might clarify the nature of participants' strategy, we analyzed infants' looking behavior by means of more fine grained measures than cumulative looking time at the outcomes. In Experiments 4 to 6 we recorded participants gazing at the stimuli during the entire experiment by means of an eye-tracker. We analyzed the dynamics of participants' attention in terms of the pattern of gaze point shifts between the elements of the scene. We also analyzed participants' shifts dynamics after the outcome (i.e. during the measurement of the cumulative looking time) and before the outcome (i.e. while they were watching the development of the movie and were likely to form the expectations that had driven their response to the outcome).

In the next sections, we will present our analyses of participants' (gaze) shifts dynamics in Experiments 4 to 6 and we will discuss its possible implications for the interpretation of participants' overall performance.

4.2. Inference Phase and Test Phase

The structure of our tasks (for Experiment 4 see section 2.1; for Experiment 5 and Experiment 6 see section 3.1) possibly gives us the opportunity to measure separately the shifting patterns that might have accompanied the execution of the logical inference and those that might have been executed in response to the outcome.

In our tasks we always presented the evidence for the disjunctive premise of the disjunctive syllogism (e.g. "*either* Object A *or* Object B is inside the cup") before the evidence for the negative premise (e.g. "Object A is *not* inside the cup"). That is, the scooping event always precedes the revelation of which object is not inside the cup (revelation made either by removing the occluder in Experiment 4 or else by having one object exit from behind the occluder in Experiments 5 and 6).

Notice that in our tasks it is only after an object has been retrieved outside the cup, and thus allows the negative premise to be set up, that what object is inside the cup can be inferred by logic. As a result, we could control the precise moment in which participants enter in the "epistemic condition" to complete the syllogism. That is, we could control the moment when the participants were presented with all of the evidence required to set up both premises of the syllogism and thus required to infer by disjunction elimination which object is inside the cup.

Afterwards, the outcomes (Consistent or Inconsistent) are used to test whether participants performed the inference. If participants worked out, or had already worked out, the conclusion of the disjunctive syllogism, they may have detected a mismatch between their rational expectations and the outcome.

We will refer to the part of a trial that follows the moment when the evidence for the negative premise is provided up to the moment when the outcome is revealed as "Inference Phase". The part of the trial following the outcome up to the end of the trial is the Test Phase (see Figure 20).

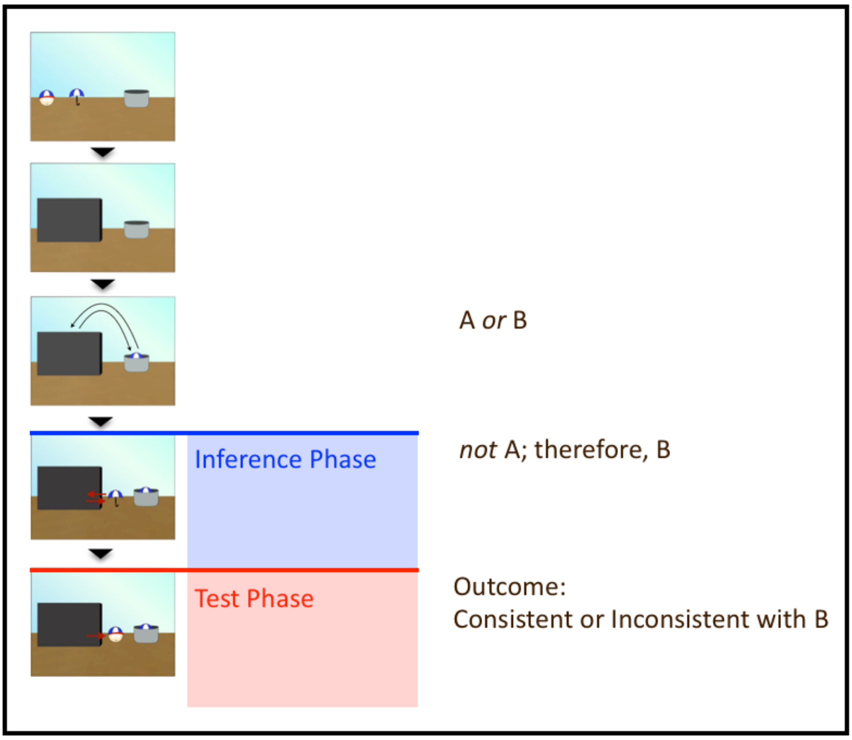

Figure 20. The Inference Phase follows the moment when the evidence for the negative premise is provided up to the moment when the outcome is revealed (in blue). The Test Phase follows the outcome up to the end of the trial (in red).

4.3. Gaze shifts dynamics in the Inference Phase: Experiment 4

In Experiment 4 (see section 2.5) we presented 12-month-old infants with the Scooped Object task and we used the eye-tracker to record participants' gaze during the entire task.

The Inference Phase of the current experiment was defined as beginning when the occluder was entirely removed and as terminating when the cup started to reveal its content, and lasted for a total of 2 s (see Figure 21).

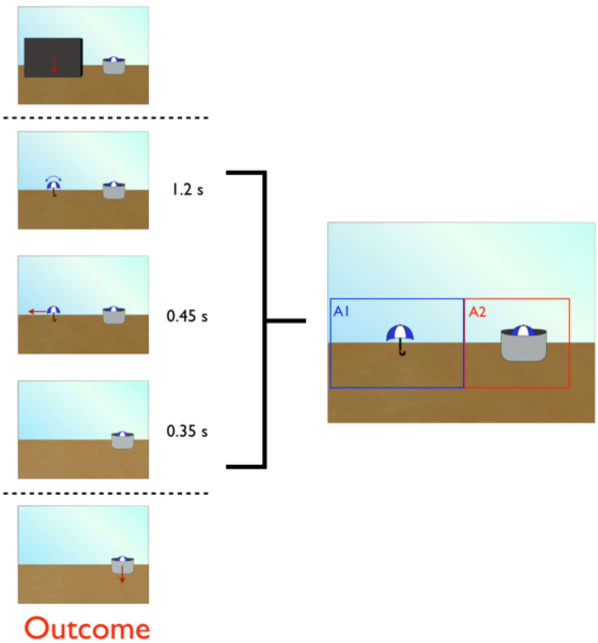

Figure 21. The Inference Phase in Experiment 4. The Inference Phase started when the object retrieved behind the occluder was fully visible for 1.2 s. Afterwards, the object moved toward the left edge of the stimuli window and completely exited from the scene after 0.45 s. After 0.35 ms from the complete disappearance of the object, the cup started to release its content and the Inference Phase terminates

It is important to clarify that nothing in our design or in the DS hypothesis implies that infants *must* complete the logical inference in the Inference Phase. It is presumably possible for a participant to work out the syllogism in the Test Phase instead. For example, it seems possible for a viewer to maintain in memory the information acquired before the Test Phase and to realize that an outcome is inconsistent by performing a disjunctive syllogism after the outcome has been produced, even several seconds after, based on information stored in memory. Thus, it is possible in principle for a participant to perform the inference either in the Inference Phase or in the Test Phase.

However, another possibility is that participants will perform the inference mostly in the Inference Phase, rather than in the Test Phase. For example a disjunctive syllogism might be a spontaneous inference that is automatically executed as soon as both its premises are represented in working memory ((Braine & O'Brien, 1998); (Lea, 1995); but see (Reverberi, Pischedda, Burigo, & Cherubini, 2012)). Alternatively, irrespective of the spontaneity of the disjunctive syllogism, infants' and toddlers' memory resources might be too limited to reliably support a postponed execution of the inference. Their memory resources might be too limited to maintain all information required for the inference until the Test Phase (e.g. that one of the two objects is inside the cup and Object A has exited from behind the occluder). While a swift execution of the inference will discharge memory load by integrating that information (e.g. that Object A has exited from behind the occluder and Object B is inside the cup).

Finally, it is possible in principle that participants will not be able to complete the disjunctive syllogism before the beginning of the Test

Phase. For example, the execution of the disjunctive syllogism might require more than the 2 s (the length of the inference phase) for infants and toddlers.

We were looking for possible gazing patterns indicative of the execution of the disjunctive syllogism during the Inference Phase. Broadly speaking, the logical inference involves both the retrieved object and the one hidden inside the cup, since the identity of one object is deduced from the identity of the other (e.g. the object inside the cup is *either* Object A *or* Object B; Object A is *not* the object inside the cup; therefore the object inside the cup is Object B). From the point of view of a rational viewer, there is a logical relation between the identities of the two objects.

Therefore, we tentatively hypothesize that the performance of the disjunctive syllogism during the Inference Phase might result in the tendency to execute gaze point shifts between the retrieved object and the object hidden inside the cup, before the cup is removed and its content revealed (i.e. before the outcome).

To test whether participants' gaze dynamics is indicative of the disjunctive syllogism we reason that, if the shifts in the Inference Phase reflect participants' strategy, then higher regularity in the execution the shifts will be positively associated with participants' performance level.

Alternatively, participants' shifts dynamics in the Inference Phase might be entirely driven by visual properties of the elements of the scene, and not by logical reasoning. Such visual properties might be object saliency (Itti, Itti, & Koch, 2001) or the perceptual similarity between the upper part of the object retrieved behind the occluder and the visible upper part of the object hidden inside the cup. In this case there seems to be no obvious reason to expect a positive association of the shift in the Inference Phase and participants' performance.

To test these predictions, we counted separately in how many trials each participant executed at least one shift from the visible object to the hidden object (from now on "VH-shift", see Figure 22) and in how many trials at least one shift from the hidden object to the visible object (from now on "HV-shift", see Figure 22).

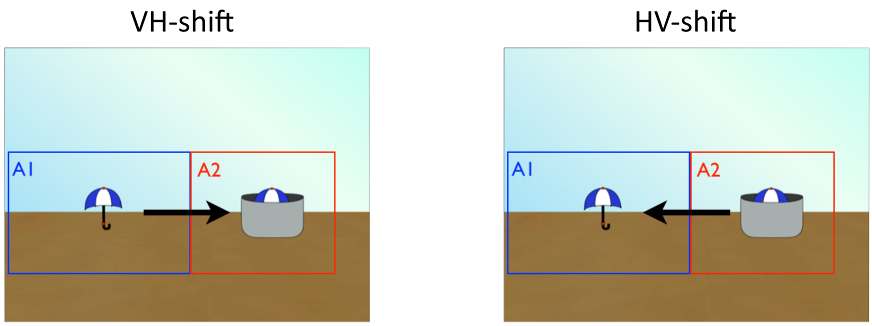

Figure 22. The VH-shift and the HV-shift in Experiment 4.

We counted one gaze point shift from one object to the other any time the gaze points of two consecutively collected samples occurred in the area associated with the first object and then in the area associated with the second object.

4.3.1. Experiment 4: results

We measured individual participant's degree of success as a difference score: average looking time at the Inconsistent Outcome minus average looking time at the Consistent Outcome.

On average, participants performed a VH-shift in the Inference Phase in 45% of the trials and a HV-shift in the Inference Phase in 27% of the trials.

In order to test whether either VH or HV-shifts might have had an impact on participants' success in detecting the Inconsistent Outcome, a separate simple linear regression analysis was performed for each type of shift.

We found that the percentage of trials in which infants performed at least one VH-shift predicted their degree of success in the task ($\beta = 3.94$; $t(21) = 2.31$, $p < 0.05$; $R^2 = 20.3\%$); (see Figure 23). In contrast, the percentage of trials in which infants performed at least one HV-shift did predict to a lesser extent participants' degree of success in the task ($\beta = 3.54$; $t(21) = 1.82$, $p = 0.083$; $R^2 = 13.6\%$).

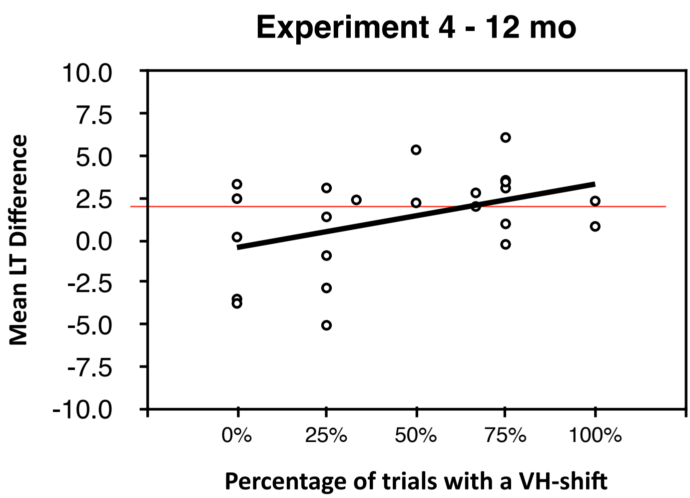

Figure 23. Scatterplot showing the relation between the percentage of trials in which infants executed at least one VH-shift and the difference between mean looking time at the Inconsistent Outcome and at the Consistent Outcome (Experiment 4).

4.3.2. Experiment 4: discussion

The analyses suggest that the regular execution of a VH-shift, and to a lesser degree of a HV-shift, during the Inference Phase was associated with better detection of the Inconsistent Outcome.

This association with participants' degree of success suggests that the shifts between the two objects in the Inference Phase, especially the VH-shift, might be indicative of the strategy behind infants' detection of the Inconsistent Outcome.

An alternative interpretation of the result can be developed starting from the observation that higher regularity in the execution of shifts between the two objects during the Inference Phase might simply reflect a higher attention and exploratory attitude toward the elements of the scene. It is also possible that such higher generic attention and exploratory attitude might have supported the detection of the Inconsistent Outcome in Experiment 4 by means of the strategy used by participants. In particular, this seems to be possible no matter whether that strategy is based on logical representations, set representations or any other short of representations. As a result, the detected association might be explained by denying that the shifts are indicative of infants' specific strategy, and by proposing that they reflect generic attention and an exploratory attitude toward the elements of the scene.

The analyses of the infants' shifts dynamics in Experiment 6, discussed in the next section, provide the opportunity to test these two competing explanations of the association.

4.4. Gaze shifts dynamics in the Inference Phase: Experiment 6

Experiment 6 makes use of the Hidden Object task (see section 3.1.). In the Hidden Object task, the occluder is never removed, the first object is retrieved when it exits from behind the occluder and the outcome consists, not in the disclosure of the cup content, but in having a second object exit from behind the occluder.

Because of the different task and stimuli used in the two experiments, the Inference Phase of Experiment 6 is somewhat different from that of Experiment 4. We were interested in measuring infants' gaze shift from the moment when the evidence for the last premise of the syllogism was available (i.e. after the first exit of one object from behind the occluder), before the outcome was produced (i.e. before the exit of the second object from behind the occluder). However, in Experiment 6 the time elapsed from the first object exit to the outcome is longer than in Experiment 4 (Experiment 6: 3.18 s; Experiment 4: 2 s). To better compare the two experiments we restricted our analyses to a 2 s period, so that it was identical to the Inference Phase of Experiment 4 (see Figure 24). The VH and HV-shifts were counted as in Experiment 4. However, the areas of interest were adapted to the different disposition of the objects (see Figure 24).

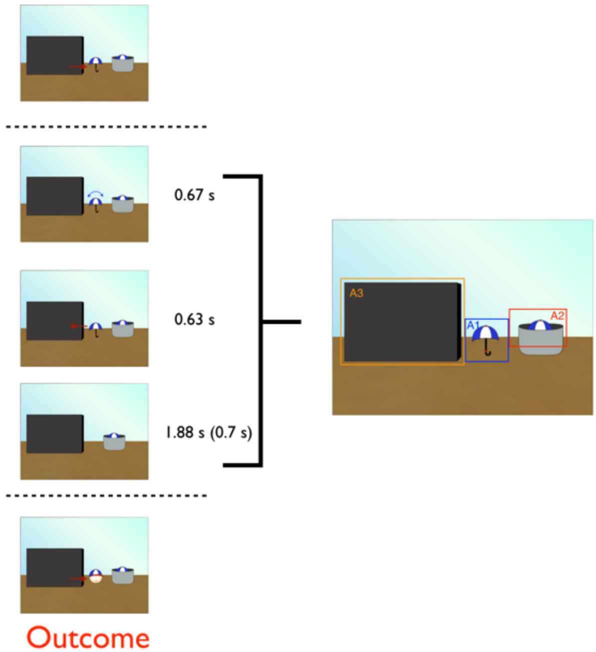

Figure 24. *The Inference Phase in Experiment 6. The Inference Phase started when the first object exited form behind the occluder and reached the position between the occluder and the cup. Afterwards the object executes its characteristic movement for about 0.67 s and then returns behind the occluder (the object has completely disappeared behind the occluder 0.63 s after the end of the characteristic movement). The Inference Phase terminates 0.7 s after the object disappeared behind the occluder.*

Experiment 6 gives the opportunity to test if the shifts in the Inference Phase reflected generic attention and an exploratory attitude toward the elements of the scene, rather than a specific strategy that infants adopted to solve the task (see section 4.3.2).

Three elements are present in the scene of the Hidden Object task during the Inference Phase: the occluder, the entirely visible object and the cup.

If the VH-shift is associated with infants' performance because it reflects generic attention and an exploratory attitude, then the gaze point shifts toward the other element of the scene, such as shifts between the visible object and the occluder, might also have a similar association with participants' success.

In contrast, if the shifts between the visible object and the object inside the cup, in particular the VH-shifts, are specifically indicative of the strategy the infants used to solve the task, then we might expect those shifts to be associated with better detection of the Inconsistent Outcome, while other types of shifts will not be, or will be only to a lesser degree, positively associated with participants' performance level.

To test these predictions, we counted separately in how many trials each participant executed at least one VH-shift (see section 4.3), in how many trials at least one HV-shift (see section 4.3), in how many trials each participant executed at least one shift from the visible object to the occluder (VO-shift) and in how many trials at least one shift from the occluder to the visible object (OV-shift) (see Figure 25).

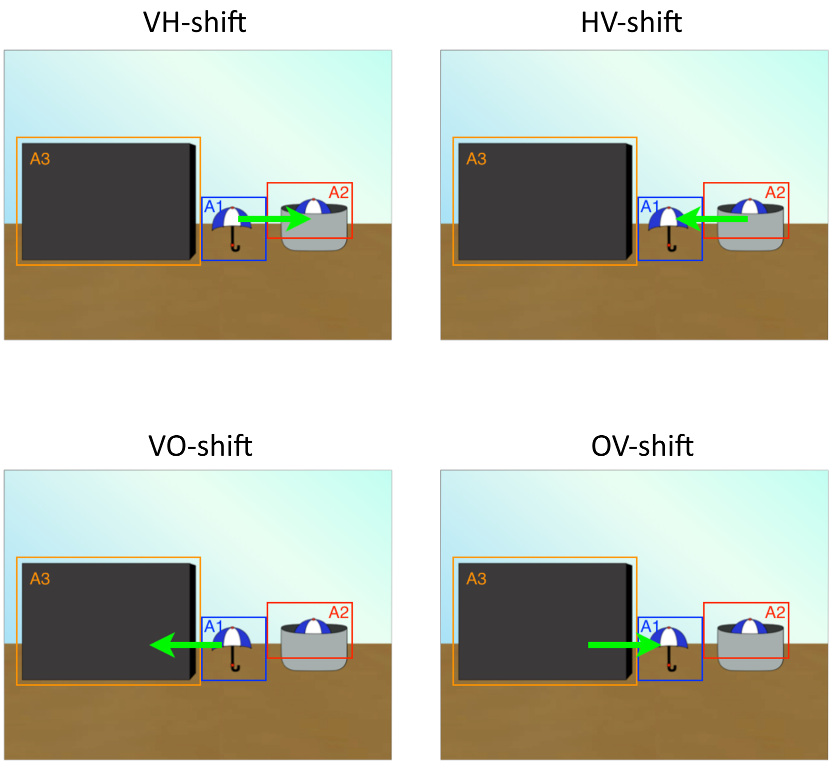

Figure 25. The VH-shift, the HV-shift, VO-shift and the OV-shift in Experiment 6.

We counted one gaze point shift from one object to the other any time the points of gaze of two samples consecutively collected by the eye-tracker were estimated to have occurred in the area associated with in the first object and in the area associated with the second object, respectively.

4.4.1. Experiment 6: results

On average, participants performed (at least) a VH-shift in 56% of the trials, an HV-shift in 18% of the trials, a VO-shift in 63% of the trials and an OV-shift in 41% of the trials.

To test whether either a VH- or HV-shift, or a VO- or OV-shift might have had an impact on participant success in detecting the Inconsistent Outcome, a separate simple linear regression analysis was performed for each type of shift.

One observation was removed from the sample as it had an externally studentized residual larger than 2.5 (in all four regressions).

The percentage of trials in which infants performed at least one VH-shift predicts their degree of success in the task ($\beta = 4.12$; $t(21) = 2.96$, $p < 0.01$; $R^2 = 29.4\%$); (see Figure 26). In contrast, participants' performance was not predicted by either the HV-shift regularity ($\beta = 1.55$; $t(21) = 0.65$, $p = 0.521$) or the VO-shift regularity ($\beta = -2.29$; $t(21) = -1.38$, $p = 0.181$) or the OV-shift regularity ($\beta = 0.28$; $t(21) = 0.13$, $p = 0.896$).

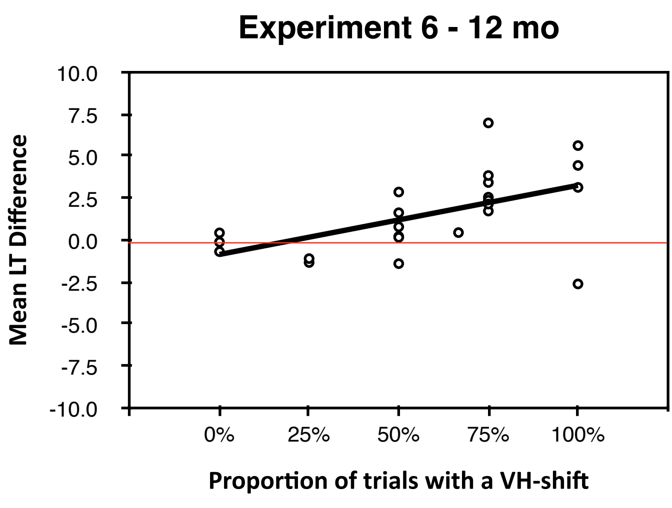

Figure 26. Scatterplot showing the relation between between the percentage of trials in which infants executed at least one VH-shift and the difference between mean looking time at the Inconsistent Outcome and at the Consistent Outcome (Experiment 6).

4.4.2. Experiment 6: discussion

The results showed that also in the Hidden Object task infants' higher degree of regularity in the VH-shift was associated with a more effective detection of the Inconsistent Outcome. This is what we would expect if the VH-shift reflected infants' strategy in the task.

Infants' behavior also provides some evidence that the VH-shift might not merely reflect generic attention and an exploratory attitude toward the elements of the scene. Although shifts between the visible object and the occluder were rather frequent during the Inference Phase, only the VH-shift toward the object hidden inside the cup was positively associated with participants' success in the task. The specificity of the association of VH-shift with better performance is not what we might expect if the fostering effect of the VH-shift is just a consequence of participants' generic attention and exploratory attitude toward the elements of the scene.

In the next section, we analyzed the toddlers' shifts dynamics in Experiment 5, to assess if their shifts dynamics in the Inference Phase was also associated with their performance in the task.

4.5. Gaze shifts dynamics in the Inference Phase: Experiment 5

In Experiment 5 we recorded 19-month-old toddlers' gazing behavior in the Hidden Object task. We analyzed toddlers' gaze shifts dynamics during the Inference Phase to test if an association with participants' performance took place, as it did for infants.

Although the test movies of Experiment 5 have the same structure as the movies of Experiment 6, the timing of some events is slightly different (see section 3.3.1.2.). As a result, the timing of the key events of the Inference Phase of Experiment 5 is also slightly different than that of Experiment 6.

Also in Experiment 5, we were interested in measuring participants' saccadic dynamics starting from the moment when the evidence for the last premise of the syllogism became available (i.e. the moment when one object exits from behind the occluder for the first time) but before the second exit from behind the occluder. However, in Experiment 5 the time elapsed from the first object being fully visible until the outcome is different from Experiment 4 and Experiment 6 (Experiment 5: 2.55 ; Experiment 6: 3.18 s; Experiment 4: 2 s). To better compare the three experiments we restricted our analyses to a 2 s period, so that they were identical to the Inference Phase of Experiment 4 (see Figure 27).

To test potential associations of toddlers' gaze point shifts dynamics with their level of performance, we coded their gazing behavior exactly as we did in Experiment 6 (see section 4.4.).

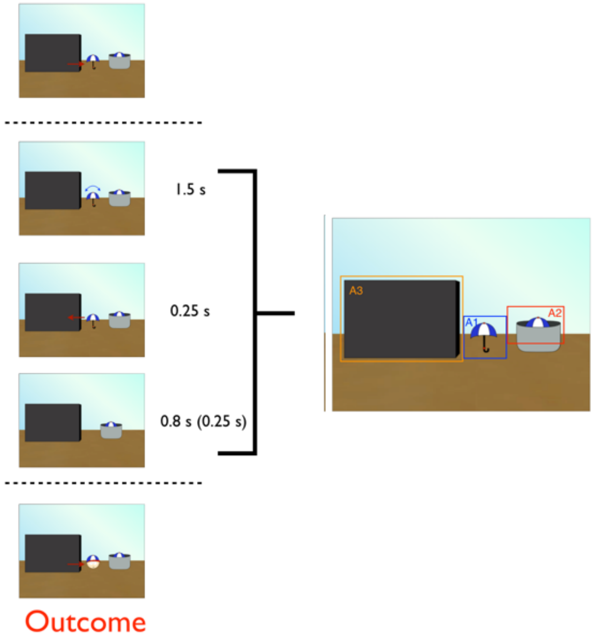

Figure 27. The Inference Phase in Experiment 5. The Inference phase of Experiment 5 started when the first object exited from behind the occluder and reached the position between the occluder and the cup. The object then executed its characteristic movement for about 1.5 s and returned behind the occluder (the object has completely disappeared behind the occluder 0.25 s after the end of the greeting movement). The Inference Phase terminates 0.25 s after the object disappeared behind the occluder.

4.5.1. Experiment 5: results

On average, participants performed (at least) a VH-shift in 56% of the trials, an HV-shift in 39% of the trials, a VO-shift in 21% of the trials and an OV-shift in 0.06% of the trials.

The percentage of trials in which toddlers performed at least one VH-shift did not predict their degree of success in the task (β = 1.14; $t(22)$ = 0.69, p = 0.51); (see Figure 28). Furthermore, participants' performance was not predicted by either the HV-shift regularity (β =

1.47; *t*(21) = 0.99, *p* = 0.333) or the VO-shift regularity (*β* = -1.01; *t*(21) = -0.53, *p* = 0.601) or the OV-shift regularity (*β* = -3.20; *t*(21) = -0.774, *p* = 0.447).

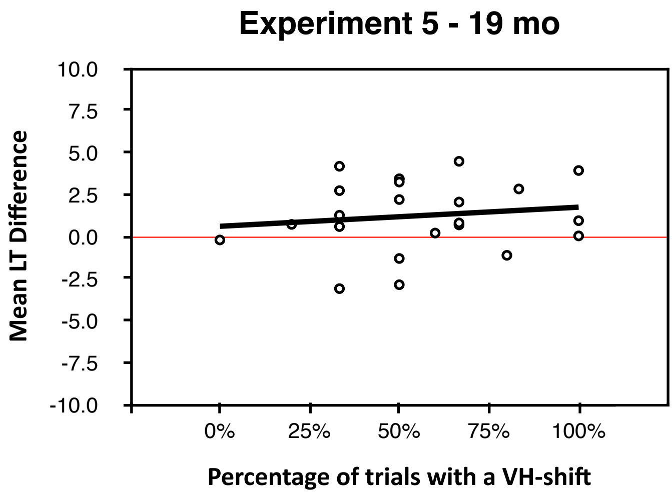

Figure 28. Scatterplot showing the relation between between the percentage of trials in which infants executed at least one VH-shift and the difference between mean looking time at the Inconsistent Outcome and at the Consistent Outcome (Experiment 5).

4.5.2. Experiment 5: discussion

Although toddlers as a group succeed in the Hidden Object task (see section 3.2.3), their degree of success was positively associated neither with the regularity in performing the VH-shift in the Inference Phase, nor with the regularity of either the HV-shift, the VO-shift or the OV-shift.

The absence of such association suggests that the VH-shift in the Inference Phase is a less efficient marker of the strategy adopted by toddlers, than of the strategy adopted by infants.

Given the preliminary character of these results we do not endorse a specific explanation of the difference between the two age groups. However, in the general discussion of the gaze dynamics in the Inference Phase (next section) we will tentatively sketch three alternative interpretations as possible guidelines for further investigation of the different behaviors of infants and toddlers.

4.6. Gaze shifts dynamics in the Inference Phase: general discussion

Analyses of 12-month-old infants' gazing behavior revealed that the regular execution of a gaze point shift from the first retrieved object to the object hidden inside the cup during the Inference Phase was associated with better performance, in both the Ambiguous Containment and the Hidden Object task.

The association between the regular execution of the shift and participants' performance level suggests that the shift might be indicative of the strategy adopted to solve the task.

Alternatively, it is possible that, although the shift is a marker of the participants with better performance, it is not related to the strategy that infants adopted to solve the task per se. Namely, it is possible that the shift merely reflects a higher generic attention and exploratory attitude (instead of a more specific strategy). Plausibly, a higher generic attention and exploratory attitude might promote the detection of the Inconsistent Outcome by means of any sort of plausible strategy, and so may explain the association of the shift with the performance level.

We think that the available evidence speaks against this possibility. In the Hidden Content task several elements were present in the scene during the Inference Phase (the occluder, the visible object, the cup). During the Inference Phase infants executed shifts toward each of those elements (not only from the retrieved object to the object hidden inside the cup). However, of the several types of shifts measured, only the shifts from the retrieved object to the object hidden inside the cup were positively associated with participants' success in the task. The specificity of the association of the VH-shift with better performance is not what we would expect if the shift merely reflects participants' generic attention and exploratory attitude toward the elements of the scene.

Thus, the strongest association of the VH-shift with the level of performance (compared to the other type of shifts executed in the Inference Phase) suggests that the VH-shift might actually be indicative of the strategy adopted by infants to solve the task. If this conclusion is correct, the character of the VH-shift may provide useful information regarding the nature of the strategy that infants adopted to solve the task. Although the available evidence might not yet be sufficient to determine the strategy infants followed, we will tentatively discuss some characteristics of the VH-shift that, we think, are preliminary evidence that their strategy might be based on a disjunctive syllogism.

First, the shift occurred after the last piece of evidence required for the logical inference was provided (i.e. evidence of which object was not inside the cup).

Secondly, the scope of the shift corresponds to the scope of the disjunctive syllogism. The shift from the visible object to the object hidden inside the cup was associated with the performance level. The shift from the visible object to the occluder was not, although it was relatively frequent in the Inference Phase. This pattern of results suggests that the decisive shift is the one involving the visible object and the one hidden inside the cup. Indeed, in the disjunctive syllogism those two objects are crucial, because the identity of one object is inferred from the identity of the other. In this sense, the scope of the shift and the scope of the syllogism seem to correspond.

Thirdly, the direction of the shift is analogous to the direction of the logical inference. The shifts that predicted infants' degree of success were precisely those from the visible object outside the cup to the object hidden inside the cup. Similarly, in the strategy proposed by the DS hypothesis the identity of the object inside the cup is inferred from the identity of the visible object outside ("Object A is outside the cup; Object A is *not* inside the cup; therefore Object B is inside the cup"). In other words, perceptual evidence relative to the visible object grounds a logical inference regarding the identity of the object inside the cup. Thus, the gaze point shift has the same direction as the conclusion of logical inference in the sense that, in both processes, the focus of the process (point of gaze for the shift and subject matter for the inference) moves from the visible to the hidden object.

We consider that the timing, scope and direction of the shift associated with infants' success suggest that the strategy behind infants' performance is based on the disjunctive syllogism.

In contrast with infants, and although toddlers did solve the Hidden Object task, toddlers' degree of success seems to not be associated with the regular execution of the VH-shift in the Inference Phase.

Although the available evidence might be insufficient to clarify the nature of the difference between infants' and toddlers' responses, we will tentatively sketch two alternative interpretations as possible guidelines for future investigation.

One explanation for the presence of the association in infants', but not in toddlers', performance might be that toddlers but not infants have the memory resources to perform a disjunctive syllogism post-hoc (in the Test Phase). It is indeed possible that memory limit leaves infants without the option of postponing the inference to the Test Phase (see section 4.3.). Thus, an infant will either complete the disjunctive syllogism in the Inference Phase as soon as the required evidence is available, or else he will risk failing in the task. As a result, infants' performance will depend on a regular execution of the disjunctive syllogism and be associated with the VH-shift in the Inference Phase, if the shift is associated with the execution of the inference.

By the same token, it is possible that toddlers' memory resources give the option of working out the logical inference post-hoc. Thus, toddlers' performance in the task depends less strongly on the execution of the inference as soon as all required information is available (in the Inference Phase), since if they miss that opportunity they can solve the logical problem during the Test Phase. As a result, toddlers' performance might not be associated with the VH-shit in the Inference Phase, although the shift is indicative of the inference.

An alternative explanation might be that although both infants and toddlers tend to perform the inference as soon as the evidence is available, either because of the automaticity of the disjunctive syllogism (see section 4.3.) or because of memory limitation, toddlers' gaze point orientation was less influenced by the drawing of the inference than the gaze point orientation of infants. Infants and toddlers might have drawn the inference at the same moment, in the Inference Phase. However, toddlers may not have to shift their gaze between the objects to do that. Therefore, toddlers' gaze point shifts might be less indicative of the nature of their strategy.

Further investigations might explore these alternative explanations. A possible starting point could be to test whether the association between shifts and successful detection of the Inconsistent Outcome is present in adults.

In summary, we propose that the specificity of the association of pre-outcome VH-shift and infants' performance level suggests that the shift might be indicative of the strategy that infants used to solve our tasks. Although further evidence is required to clarify the nature of infants' strategy, we think that the timing, scope and direction of the shift are suggestive of a strategy being based on the disjunctive syllogism. The association was absent in toddlers' responses. The difference between the two age groups might reflect toddlers' richer resources. Further investigations are required to clarify the nature of this difference.

4.7. Gaze shifts dynamics in the Test Phase: Experiments 5 and 6

In this chapter we report the analyses of participants' shifts dynamics in response to the outcomes of our task (i.e. their shift patterns in the Test Phase, see section 4.3.). The looking time results suggest that both infants and toddlers were able to solve our tasks, since both groups looked at the Inconsistent Outcome for longer than at the Consistent one. However, the looking time response does not entirely clarify which strategy was behind the detection of the Inconsistent Outcome (see sections 4.1.). Measuring the type and amount of gaze point shifts executed while looking at the outcomes might offer a more fine grained characterization of participants' looking response and help to clarify the nature of participants' strategy in the task.

We focused our analyses on the shifts dynamics in the Test Phase of Experiments 5 and 6. In those experiments, the 19- and 12-month-olds, respectively, were presented with the Hidden Object task (see section 3.1.). We focused on the Hidden Object task as multiple objects were present in the scene in its outcome. A display with multiple objects might foster the execution of exploratory visual pattern possibly reflecting participants' strategies in processing the display. We did not analyze infants' shift responses in Experiment 4, since in the Ambiguous Containment only one object is present in the scene (see section 2.1.), and thus no gaze shifts can be executed between objects in the scene.

In the Hidden Object task, participants witnessed an object exit from behind the Occluder for a second time. The Test Phase began when the object reached the middle point between the occluder and the cup. Once the Test Phase had begun the movie was paused. As a result, the stimuli of the Test Phase were static pictures (see Figure 29).

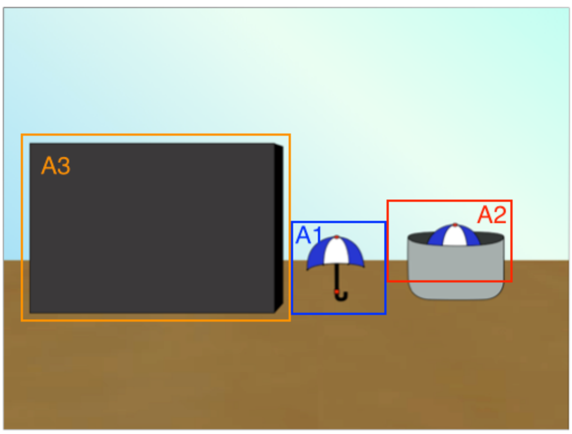

Figure 29. Outcome of the Hidden Object task with its three areas of interest.

It is a plausible working hypothesis that when participants were presented with the outcomes, they might have tried to identify the objects in the outcome, based on previously acquired expectations (Leslie et al., 1998; Fei Xu, 2007). In other words, presented with the outcome, participants might have compared the objects perceived in the scene with a previously formed internal representation of the objects in the scene.

The DS hypothesis and the alternative explanations individuated in the discussion of the looking time results (see section 2.6. and 3.4.3.), propose that the detected looking time difference is a result of the detection of the incongruence of the Inconsistent Outcome with participants' expectations.

To better characterize the nature of such a mismatch we tested whether participants' shifts dynamics in the Test Phase, the shifts that were executed between the elements of the scene, were modulated by the Outcome Type (Consistent vs. Inconsistent).

If the shifts dynamics in the Test Phase reflects processes that are related to the detection of the mismatch in the Inconsistent Outcome, the shifts dynamics might be modulated by the outcome type.

We also think that the DS hypothesis might motivate a more specific prediction regarding the way the Inconsistent Outcome might modulate participants' shifts dynamics.

According to the DS hypothesis, the detection of the Inconsistent Outcome is due to participants' conclusions that Object B has to be inside the cup (since Object A was found outside the cup). In the Consistent Outcome, the object that exited from behind the occluder looked identical to Object A. In the Inconsistent Outcome, that object looked identical to Object B, the object that by logic had to be inside the cup. While the look of the visible object suggested that the visible object is Object B, by logic Object B has to be the one inside the cup. It is the mismatch of identity of these two objects that complicates the mapping of participants' representation onto the Inconsistent Outcome.

Since according to the DS hypothesis the core of the mismatch is the identity of the visible object and the identity of the object inside the cup, the hypothesis suggests that the Outcome Type should especially modulate the shifts between the visible object and the object hidden inside the cup.

We counted separately the shift between adjacent objects. That is, we counted the shifts between the visible object and the object hidden inside the cup (from now on "VorH-shift") and we counted the shifts between the visible object and the occluder (from now on "VorO-shift"). In counting the shifts we abstracted from the direction of the shift (e.g. whether VorH-shift was toward the visible object or toward the hidden object), since we had no specific prediction regarding the shifts' direction in the Test Phase (See Figure 30).

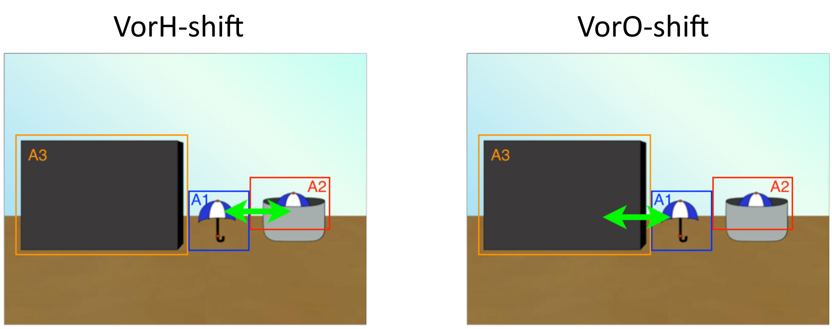

Figure 30. The VorH-shift and VorO-shift in Experiment 5 and 6.

We counted one shift between the two objects if the points of gaze of two samples consecutively collected by the eye-tracker were estimated to have occurred in the area associated with one object and in the area associated with the other object.

To compare the shifts dynamics in response to the different types of outcomes, we calculated the shift-rate in a trial as the number of shifts performed in that trial divided by the looking time in that trial. We used the shift-rate, instead of simply using the number of shifts performed in a trial, to adjust our estimation of participants' likelihood to perform a shift from the effect of the trial length on the number of shifts performed. Notice that in our design the length of the Test Phase was not fixed, but contingent on the participants' looking time, as the Test Phase ended when a participant looked away for 2 consecutive seconds. Furthermore, our analyses revealed that the Inconsistent Outcome elicited longer looking time than the Consistent one (see section 3.2.2. and 3.3.3.).

If the shifts dynamics in the Test Phase reflects the processes that are related to the detection of the mismatch in the Inconsistent Outcome, the shift-rate will be modulated by the Outcome Type.

Moreover, if the representation behind the mismatch detection is that proposed by the DS hypothesis, then the outcome type modulation of the VorO-shift-rate should be particularly strong.

4.7.1. Gaze shifts dynamics in the Test Phase: results

Toddlers' and infants' shift-rate scores by trial were collapsed by Outcome Type (Consistent or Inconsistent) by calculating a by-participant mean for each of the two conditions. For each age group and for each type of shifts (VorH or VorO), one separate ANOVA with Outcome Type as a within-participants factor and Participant as random variable was performed on mean shift-rate scores.

The analyses of toddlers' VorH-shift-rate detected a main effect of the Outcome Type ($M_{Consistent}$ = 0.41 sac/s , $M_{Inconsistent}$ = 0.59 sac/s; $F(1,23) = 4.81$, $P < 0.05$), showing that 19-month-olds performed VorH-shifts at a higher rate in the Inconsistent Outcome than in the Consistent Outcome. In contrast, the analyses of infants' VorH-shift-rate did not detect a main effect of the Outcome Type ($M_{Consistent}$ = 0.46 sac/s, $M_{Inconsistent}$ = 0.58 sac/s; $F(1,23) = 1.74$, $P = 0.200$); (see Figure 31).

The analyses of VorO-shift-rate detected no main effect of Outcome Type, neither in the toddlers ($M_{Consistent}$ = 0.30 sac/s, $M_{Inconsistent}$ = 0.39 sac/s; $F(1,23) = 2.42$, $P = 0.133$), nor in the infants ($M_{Consistent}$ = 0.25 sac/s, $M_{Inconsistent}$ = 0.27 sac/s; $F(1,23) = 0.22$, $P = 0.640$).

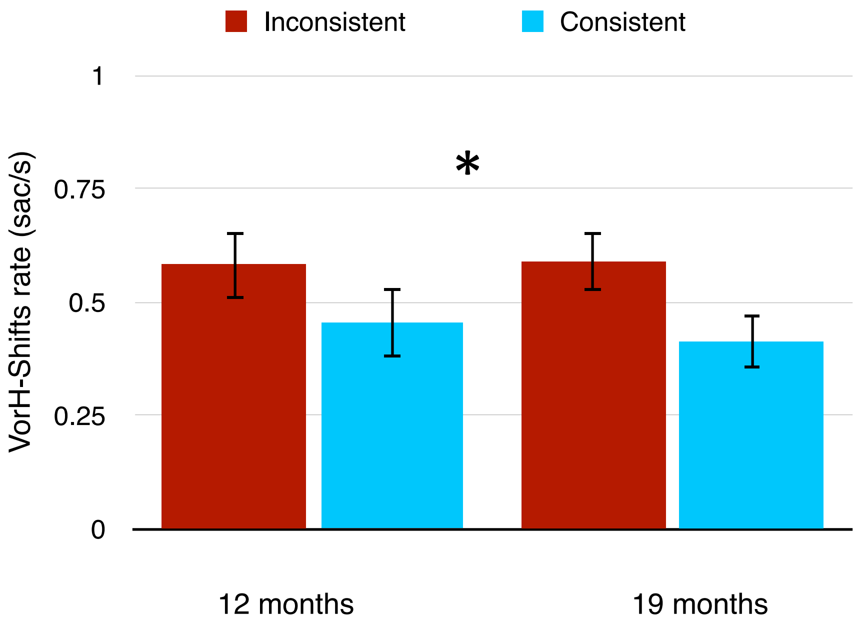

Figure 31. Mean VorH-shift-rate at the Inconsistent Outcome and Consistent Outcome in 12- and 9-month-olds'. Error bars represent SEs.

Finally, an ANOVA with Outcome Type as a within-participants factor, Age as a between-participants factor and Participant as random variable was performed on mean VorH-shift-rate scores to compare the performance of the 12-month-old infants with that of the 19-month-old toddlers. There was a main effect of Outcome Type ($M_{Consistent}$ = 0.43 sac/s, $M_{Inconsistent}$ = 0.55 sac/s; $F(1, 45) = 5.35$, $P < 0.05$) and no main effect of Age ($M_{12\text{-}months}$ = 0.49 sac/s, $M_{19\text{-}months}$ = 0.48 sac/s; $F(1, 45) = 0.02$, $P = 0.89$). Crucially, no interaction between Outcome Type and Age was detected ($F(1, 45) = 1.06$, $P = 0.31$). This pattern of results suggests that the main effect of the Outcome Type was not strongly modulated by participants' age.

4.7.2 Gaze shifts dynamics in the Test Phase: discussion

Nineteen-month-olds reacted to the Inconsistent Outcome with a higher rate of shifts between the visible object and the object hidden inside the cup, compared to the Consistent Outcome.

The infants' response was less clear: although an ANOVA focusing exclusively on the infants' performance did not detect an effect of the Outcome Type on the VorH-shifts, a between age group ANOVA detected a main effect of the Outcome Type and no interaction effect of Outcome Type and the participants age group.

While toddlers responded to the Inconsistent Outcome with a higher rate of VorH-shifts, infants possibly responded with a weaker modulation. A possible explanation for the weaker looking response of the 12-month-olds might be the development of endogenous visual attention. The detection of the Inconsistent Outcome might result in a weaker increase of shift-rate at 12 months of age than at 19 months because of the different level of development of endogenous visual attention (Colombo, 2001)

The effect of the Inconsistent Outcome on the shift-rate suggests that the shifts dynamics of toddlers, and possibly of infants, might have reflected processes that are related to the detection of the Inconsistent Outcome.

In contrast to the VorH-shifts, the rate of the shifts between the visible object and the occluder were not modulated by the Inconsistent Outcome, for neither infants nor toddlers. Thus, the effect of the Outcome Type was focused on the VorH-shifts.

If the participants' shifts dynamics reflects processes related to the detection of the Inconsistent Outcome, the focus of the modulation of the shifts dynamics might help to clarify the nature of those processes. Thus, the results suggest that the processes related to the detection of the Impossible Outcome might have been particularly focused on the visible object and the object hidden inside the cup, rather than on the occluder and the object hidden behind it.

The DS hypothesis suggests that the core of the mismatch between participants' expectation and the Inconsistent Outcome is the identity of the visible object and the hidden object.

It is worth remarking that the scope of shift-type that is modulated by the Outcome may be simply explained if the strategy behind infants' performance was based on the disjunctive syllogism.

5. GENERAL DISCUSSION

5.1. Cumulative looking time response

In a series of experiments we presented 19-month-old toddlers and 12-month-old infants with two tasks, the Scooped Object task and the Hidden Object task (see Figure 32). The tasks are designed to test their ability to infer the identity of a hidden object based on the identity of another object, by using the logical inferential rule Disjunctive Syllogism.

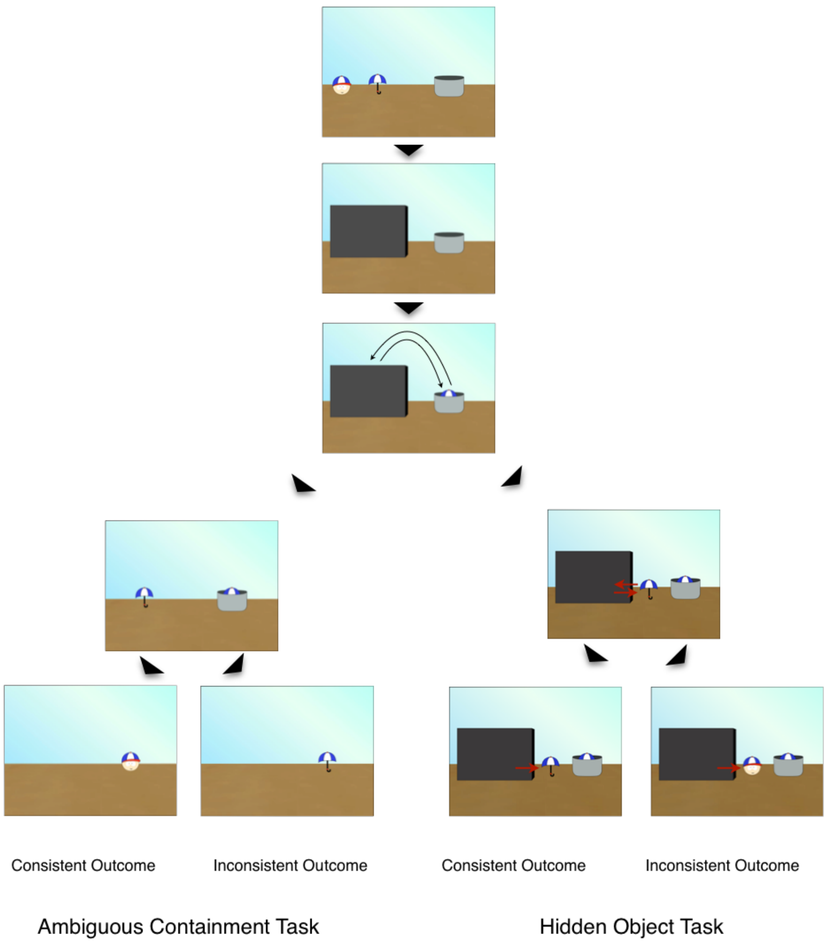

Figure 32. Schema of the Scooped Object task and the Hidden Object task.

The results of Experiments 1 to 4 suggest that 19-month-old toddlers and, under certain circumstances, 12-month-old infants reacted with surprise to the Inconsistent Outcome of the Scooped Object task. This result was predicted by the Disjunctive Syllogism hypothesis, according to which when a participant saw which object was outside of the cup, she was able to infer by disjunctive syllogism which object was inside the cup.

An alternative account of the solution of the Scooped Object task proposes that infants did not infer which object was inside the cup, but recognized the visible object based on its category. Participants may have solved the task by creating a memory model of the objects initially present in the scene and their categories (Carey & Xu, 2001; Fei Xu, 2007; J. Zosh & Feigenson, 2009). It is only once the objects became fully visible that they may have tried to map the model onto the visible objects based on their category (we will call such an account "Visible Object Mapping", henceforth VOM). For example, in the Consistent Outcome, they may have identified the face revealed by the occluder as the face initially seen and the umbrella revealed by the cup as the umbrella initially seen. Crucially, VOM accounts for surprise felt at the Inconsistent Outcome. For example, a model with just one face-object is not mapped on two distinct (serially) visible faces. The mismatch caused participants' surprise.

To test whether infants and toddlers are able to solve a similar problem when VOM is of no help, we tested them with the Hidden Object task. In Experiments 5 to 6, participants of both age groups reacted with surprise to the Inconsistent Outcome, a result that was predicted by the Disjunctive Syllogism hypothesis.

Against VOM, the result shows that participant surprise was not driven by the identification of the entirely visible objects. Indeed, in the Inconsistent Outcome condition of the Hidden Object task a face and an umbrella were (serially) visible, a sequence that is compatible with a memory model consisting of one umbrella-object and one face-object.

An alternative account of the response to the Hidden Object task proposes that participants did not infer which object was inside the cup, but count both visible objects and partially visible objects in the scene. Participants may have solved the task by creating a memory model of the objects initially present in the scene and their categories. Participants may have then counted the objects that moved out from the occluder, including the object partially visible inside the cup. Finally, they may have compared the numerosity of the set of the objects in the memory model (i.e. 2) with the numerosity of the set of objects that have been serially seen in the scene, holding the expectation that the two numbers have to coincide (Feigenson & Carey, 2003) (we will call this second account "Numerosity Comparison", henceforth NC). In the Consistent Outcome of the Hidden Object task, participants saw that two objects move out from the occluder: the object inside the cup and another object that moved outside and inside the occluder several times (e.g. one face). This is the same number as the number of objects in the memory model. Crucially, NC accounts for surprise at the Inconsistent Outcome of the Hidden Object task. In the Inconsistent Outcome, participants saw that three objects moved out from the occluder: the object inside the cup and two objects that (serially) move outside from the occluder (e.g. 1 face and 1 umbrella). This is not the same number as the number of objects in the model and this might trigger participants' surprise.

In other words, the NC account proposes that rather than infer the identity of the object inside the cup by logical reasoning, participants did not disambiguate the identity of that object; instead, they solved the task by comparing the numerosities of two sets of objects, the pair of objects seen before the occluder rose and the objects that then exited from behind the occluder. Indeed, studies on infants' numerical cognition suggest that infants as young as 14 have the ability to compare the numerosity of a sets of objects hidden in a location and the numerosity of a set of objects successively retrieved in the same location (Carey, 2009; Feigenson & Carey, 2003, 2005; Feigenson et al., 2004; J. Zosh & Feigenson, 2009).

5.2. An argument against the numerosities comparison account

The NC account is grounded in an ability that infants and toddlers are very likely to possess: the ability to compare the exact numerosities of small sets of objects. Nonetheless we think that it is far from clear that NC is adequate to explain higher surprise to the Inconsistent Outcome of the Hidden Object task. The problem with NC can be clarified by a review of infants' and toddlers' ability to represent and compare sets of objects occluded from their view.

Studies on infants' numerical cognition report that infants' and toddlers' ability to represent sets of objects is sophisticated and takes category membership into account. For example, when infants see that two identical toy balls are hidden in an occluded location, they search less in that location after that two balls have been retrieved in that location than after that only one has (Feigenson & Carey, 2003). When infants see that two toy cars and two toy cats are hidden in an occluded location, they search less in that location after that two toy car and two toy cat have been retrieved in than location, than after that only one toy car and one toy cat have (Feigenson & Halberda, 2008). When toddlers see that one car and one shoe is hidden in an occluded location, they search less in that location after that one toy car and one toy shoe has been retrieved in that location, than after that one toy car and one toy cat have (J. Zosh & Feigenson, 2009; J. M. Zosh & Feigenson, 2012).

These results are interpreted as suggesting that when toddlers, and possibly infants, see few objects hidden in an occluded location they expect that the same number of objects *of each category* is hidden in that location. And that they can use their ability to represent and compare sets to check if the numbers add up, especially the ability to sort objects in sets based on their category membership. For example, if a toddler sees that one car and one shoe are hidden in an occluded location and one cat and one shoe are taken out from the location, he will expect that a car is still hidden in that location, and will keep searching for the car (J. Zosh & Feigenson, 2009; J. M. Zosh & Feigenson, 2012)

The attribution of such expectations and such representational abilities, although very sophisticated, do not explain participants' response to the Hidden Object task.

In the Hidden Object task, participants saw two objects of different categories, for example one face and one umbrella, get hidden behind an occluder (see Figure 33).

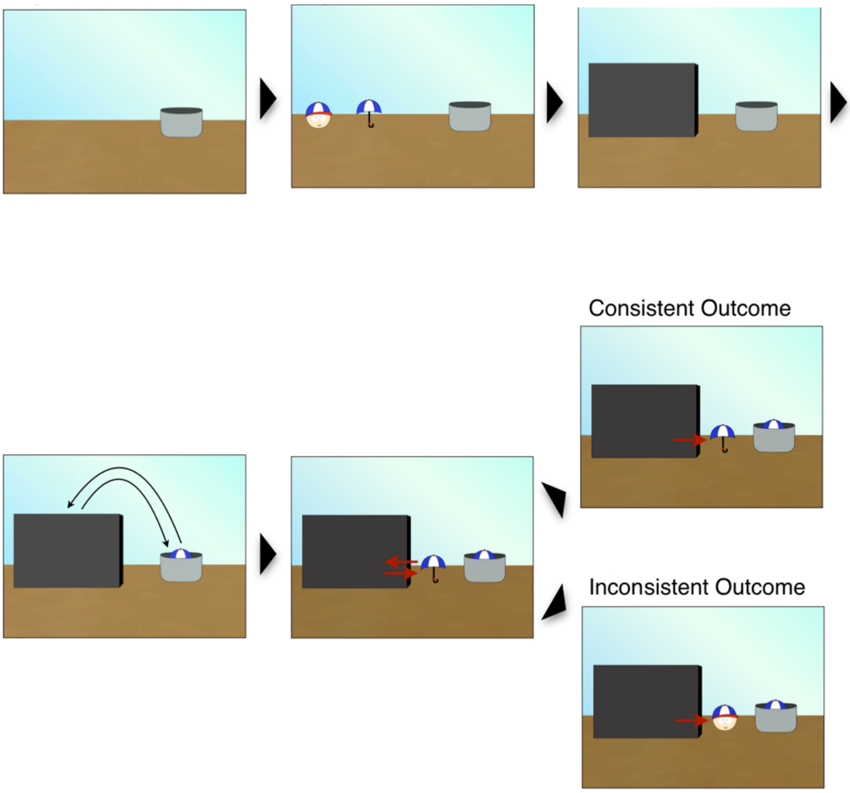

Figure 33. Schema of the Hidden Object task.

In the Consistent Outcome condition, participants saw that, one after the other, one mysterious hidden object (hidden inside the cup) and one visible umbrella came out from the occluder (the umbrella moves out, back and then out again from the occluder). How many objects might participants expect to be present in the scene?

They should expect that one face is behind the occluder, for they know that behind the occluder there were one face and one umbrella, and they saw one umbrella but no face move out from the occluder (i.e. by assumption they do not know that inside the cup there is a face). But if participants expected that the face is still behind the occluder, then the numbers in the Consistent Outcome do not add up.

There was an extra object in the scene: one face, one umbrella and one mysterious object hidden inside the cup.

In the Inconsistent Outcome, participants saw that, one after the other, one mysterious hidden object (hidden inside the cup), one visible umbrella and one visible face came out from the occluder. Also in this case the numbers did not add up. There was an extra object in the scene: one face, one umbrella and a mysterious object hidden inside the cup.

But if an extra object is detected in both types of outcome, participants should not have been more surprised by the Inconsistent Outcome than by the Consistent Outcome as result of a sets comparison. In fact, this was not the response that infants and toddlers gave. They did find the Inconsistent Outcome more surprising than the Consent one.

Briefly, in the Inconsistent Outcome, participants saw that one face and one umbrella moved out from behind the occluder. In the Consistent Outcome participants saw that one umbrella but no face moved out from behind the occluder (by assumption they did not infer that there is a face inside the cup), thus they should have expected that the face was still behind the Occluder. Therefore, in both outcomes the numbers did not add up, because a third mysterious object was hidden inside the cup. As a result, a comparison between the number of visible objects in each category does not explain the higher surprise for the Inconsistent Outcome.

In summary, recent studies suggest that toddlers' and, possibly, infants' ability to compare the numerosities of small sets of objects responds to objects' categories. It is then not clear how to explain the higher surprise of the Inconsistent Outcome in terms of sets numerosities comparison, since in both outcome types the number of objects per category does not add up, unless the category of the hidden object is inferred.

5.3. Eye-movements and participants' inferences

In three experiments we used an eye tracker to record the gazing pattern executed by participants while watching the movies and in response to the outcomes (Experiment 4: 12-month-olds, Scooped Object task; Experiment 5: 19-month-olds, Hidden Object task; Experiment 6: 12-month-olds, Hidden Object task).

In both Experiment 4 and Experiment 6 we found that a specific gazing pattern executed while watching the movie was associated with infants' performance level in the task. In the movies of both the Scooped Object and the Hidden Object task, the cup moved out from the occluder with one object inside and afterwards one of the two initial objects was found to be outside the cup. In response to finding the first object infants executed a gaze shift from the visible object toward the cup. In both task types, a systematic execution of that type of shift was positively associated with degree of success. Such an association was instead absent or weaker for other types of shifts executed in the same phase of the task, a pattern in conflict with an account in terms of participants' generic attention toward the stimuli.

The positive association with degree of success is explained if the shift is indicative of the strategy adopted by the 12-month-olds to solve the tasks. What does the shift tell us of infants' strategy?

According to the DS hypothesis, infants used evidence regarding which object is not inside the cup to infer which object is inside by disjunctive syllogism. The systematic execution of such an inference, in response to finding which object is outside the cup, would hence promote the solution of the task. We propose that the execution of the gaze shift from the visible object to the cup in response to finding which object is outside the cup is associated with the performance level because it is indicative of the execution of the logical inference.

In other words, consider three properties of the shift. First, the shift happened after the finding of the first object and before the outcome. Secondly, the shift is from the visible object to the cup. Thirdly, the shift is positively associated with performance level. These three properties of the shift are explained if the gaze shift is indicative of a disjunctive syllogism. First, the shift happened after the finding of the first object, because finding the first object is required to trigger the inference. Secondly, the shift is from the visible object to the cup, because, in the inference, evidence of which object is outside the cup is used to infer which object is inside the cup. Thirdly, the shift is positively associated with performance, because having inferred which object is inside the cup causes higher surprise to the Inconsistent Outcome.

In summary, we think that the association detected between that particular pre-outcome gaze pattern and performance offers preliminary evidence that infants might have solved our task by means of a disjunctive syllogism. However, additional evidence is required to strengthen this result. Future experiments might be specifically designed to test the hypothesis that the infants' gaze pattern reflects a logical inference.

In contrast with infants, the association has not been detected in toddlers. We tentatively proposed two possible accounts of this difference to be tested in future experiments.

A first possibility is that toddlers' higher memory resources made the execution of the disjunctive syllogism before the outcome less indispensable for solving the task. If a toddler happened to not perform the disjunctive syllogism before the outcome, he might be able to easily do that after the outcome based on information stored in memory. In contrast, infants may lack the memory resources to rely effectively on a post-outcome disjunctive syllogism. As a result, the pre-outcome disjunctive syllogism has a stronger impact on infants' performance than on toddlers' performance. Thus, it is possible that, although in both age groups the shift is indicative of the syllogism, the shift is only correlated with success in infants.

A second possibility is that, in contrast with infants, toddlers' shift in the pre-outcome phase does not reflect the execution of the disjunctive syllogism, although a disjunctive syllogism was executed in the Inference Phase both by infants and toddlers. This might be the case either because toddlers do not need to shift their gaze from the visible object to move their attention to the one hidden inside the cup, or because toddlers have a higher control of their gaze and can think of the content of the cup while fixating on the visible object. In either case, toddlers' gaze point shift might be less indicative of the nature of their strategy.

6. CONCLUSION

As we have seen (see Preface), radically different accounts of the ontogenesis of logical concepts have been proposed. According to Piaget, propositional reasoning is the product and peak of cognitive development (Piaget, 1953, 1960), Fodor's famous nativist argument entails that logical concepts are (very likely) not learned at all (Fodor, 1975).

One source of evidence of logical representations is overt linguistic behavior. The attested onset of the early logical use and understanding of logic-related words is successive to the second birthday: before that age, production and comprehension of those words provide no evidence of logical representation. Of course this is not sufficient to show that logical reasoning is absent before the age of two. Thus, the possibility of the existence of preverbal logical representations and their development in the first years of life demand alternative forms of empirical investigation.

Potential cues of preverbal logical reasoning can be found in word learning. Infants' ability to disambiguate the referent of novel names has been associated with the logically valid inferential rule disjunctive syllogism. Such ability is quite robustly attested in toddlers as young as 17 months (Halberda, 2003; Houston-Price et al., 2010; Mather & Plunkett, 2011) and occasionally documented in younger infants (Markman et al., 2003; Fei Xu et al., 2005). However, alternative accounts of novel names disambiguation have been proposed.

In contrast, success in non-verbal tasks designed to test disjunctive syllogism has never been documented in children younger than 2.5 years (Call & Carpenter, 2000; Hill et al., 2012; Premack & Premack, 1994; Watson et al., 2001). Furthermore, the evidential support that these types of tasks actually offer to the attribution of logical inferences has been questioned (Penn & Povinelli, 2007).

Thus, notwithstanding their non-conclusive character, there are cues of logical reasoning in novel name referent disambiguation in the second year of life but none in tasks of the non-verbal type. This asymmetry is suggestive of the possibility that, before the second year of life, logical inferences are uniquely at the service of language processing or learning. Alternatively, the lack of cues of preverbal logical reasoning outside of language processing might be an artifact of the lack of adequate procedure for testing it.

As result of this discussion we individuated two questions from which an investigation of the origin of logical concepts may commence: whether there are preverbal logical concepts at all and whether preverbal logical reasoning is ever employed outside of word learning and processing. The aim of this thesis was therefore to explore preverbal logical reasoning outside of the contest of language learning and processing.

To answer the two aforementioned questions together, we developed a non-verbal procedure to test infants' and young toddlers' ability to execute the disjunctive syllogism, a logical inference. In our tasks a logical problem is presented with entirely non-linguistic stimuli and participants' responses are obtained by recording and analyzing different aspects of their looking behavior.

We tested 19-month-old toddlers and 12-month-old infants with our tasks. The toddlers were of an age at which novel names referent disambiguation is attested quite robustly, the infants of the earliest age at which that ability has ever been attested. Both ages precede the attested onset of logical use and understanding of logic-related words.

The first important finding of our research is that both infants and toddlers respond with an increase of surprise to situations that violate the conclusion of the disjunctive syllogism. Indeed, analyses of cumulative looking time in six experiments revealed that the 19-month-olds' response was in accordance with the use of the disjunctive syllogism. The 12-month-olds' responses were compatible with the logical inference too, once the non-logical difficulty factors of the task were mitigated and when strong object categories contrasts helped the infants to process the movies.

This result is the first evidence that infants and toddlers in their second year of life might deploy logical reasoning in tasks that are not related to language processing. First, the result offers evidence that early logical representations might predate the mastery of logic-related words. Secondly, it suggests that preverbal logical reasoning is not uniquely in the service of language processing, since it might be employed to deal with other problems, unrelated to language.

This interpretation of our finding needs to gain additional empirical support, since alternative interpretations of our results, based on infants' non-logical numerical abilities, can be proposed. We offered theoretical considerations that may question the cogency of those alternative interpretations. However, additional empirical evidence and further research are required to clarify the nature of the strategy and of the abilities behind participants' response to our tasks.

We took the first step toward the clarification of the representation and processing behind the looking time response by analyzing participants' looking behavior beyond it. We explored participants' looking response along two additional dimensions: eye-movements executed during the presentation of our movies and eye-movements executed in response to the movies' conclusions, during the measurement of cumulative looking time.

Analyses of eye-movements executed on-line during the movie presentation detected a specific pattern that was positively associated with infants' degree of success in the task. We argued that qualitative properties of such a pattern are suggestive of a strategy based on the disjunctive syllogism. In contrast to the infants, the eye-movements pattern was present in toddlers as well, but it was not associated with performance level. Although this difference requires further investigation, we tentatively propose that it might depend on toddlers' higher memory and attentional resources that made their reasoning strategy more flexible and their eye-movements less transparent of their reasoning.

Thus the second finding of this dissertation is that the analyses of on-line eye-movements provide an additional evidence that participants' response to our tasks might be grounded in a disjunctive syllogism. More broadly, this second result indicates the possibility of using analysis of eye-movement in on-line movies processing as a strategy to study logical inference in non-linguistic tasks, and track their developmental trajectory. Further studies are required to explore this possibility and to clarify the difference between the response of infants and toddlers.

Finally, analyses of eye-movements during the measurement of the cumulative looking time detected a higher rate of exploratory eye-movements between the elements of the scene, as a response to outcomes that violate the conclusion of the disjunctive syllogism. although the effect was clearer in 19-month-olds than in 12-month-olds. Thus, the ability to perform a disjunctive syllogism can explain not only modulations of the cumulative looking time but also the modulation of the rate of eye-movements.

Together, these results provide a collection of preliminary evidence of the existence of preverbal logical reasoning, since the age of the participants in our studies predate the attested onset of the mastery of logical words. The non-linguistic character of our tasks also suggests that preverbal logical reasoning might be at least partially independent from language processing and learning: infants that have not yet mastered logical words might be able to use a disjunctive syllogism to solve a task with no linguistic stimuli involved.

We think that these first exploratory results indicate the need for an exhaustive investigation of preverbal logical abilities. Such an enterprise might offer a more complete understanding of the development of mature logical representation and a clarification of the relationship between human reasoning and language.

References

Austin, K., Theakston, A., Lieven, E., & Tomasello, M. (2014). Developmental Psychology Young Children ' s Understanding of Denial Young Children ' s Understanding of Denial.

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, *20*(3), 191–208.

Bion, R. a H., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30months. *Cognition*, *126*(1), 39–53.

Block, N. (1986). Advertisement for a Semantics for Psychology. *Midwest Studies in Philosophy*, *X*, 615–678.

Bonatti, L.L., Frot, E., Zangl, R., & Mehler, J. (2002). The human first hypothesis: identification of conspecifics and individuation of objects in the young infant. *Cognitive Psychology*, *44*(4), 388–426.

Bonatti, L. L., Frot, E., & Mehler, J. (2005). What face inversion does to infants' counting abilities. *Psychological Science*, *16*(7), 506–510.

Braine, M. D. S. ., & O'Brien, D. P. (1998). *Mental logic*. Mahwah, NJ, US: Lawrence Erlbaum Associates.

Call, J., & Carpenter, M. (2000). Do apes and children know what they have seen? *Animal Cognition*, *3*(4), 207–220.

Carey, S. (1978). The child as word learner. In *Linguistic theory and psychological reality* (pp. 264–293).

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

Carey, S. (2014). On Learning New Primitives in the Language of Thought: Reply to Rey. *Mind and Language*, *29*(2), 133–166.

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: Beyond object files and object tracking. *Cognition*, *80*(1-2), 179–213.

Chomsky, N. (1959). A Review of B. F . Skinner's Verbal Behavior. *Language*, *35*(1), 26–58.

Clark, E. V. (1988). On the logic of contrast. *Journal of Child Language*, *15*(2), 317–335.

Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language*, *17*(2), 417–431.

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*, 337–67.

Crain, S., & Khlentzos, D. (2010). The logic instinct. *Mind and Language*, *25*(1), 30–65.

De Marchena, A., Eigsti, I. M., Worek, A., Ono, K. E., & Snedeker, J. (2011). Mutual exclusivity in autism spectrum disorders: Testing the pragmatic hypothesis. *Cognition*, *119*(1), 96–113.

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: a pragmatic account. *Developmental Psychology*, *37*(5), 630–641.

Dimroth, C. (2010). The acquisition of negation. In L. R. Horn (Ed.), *The Expression of Negation* (pp. 39–72). Berlin & New York: Mouton de Gruyter.

Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: Evidence from infants' manual search. *Developmental Science*, *6*(5), 568–584.

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

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314.

Feigenson, L., & Halberda, J. (2008). Conceptual knowledge increases infants' memory capacity. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(29), 9926–9930.

Fodor, J. (1975). *The language of thought*. (J. J. Katz, D. T. Langendoen, & G. A. Miller, Eds.)*New York* (Vol. 4). Harvard University Press.

Fodor, J. (2004). Having concepts: a brief refuation of the twentieth century. *Mind & Language*, *19*(1), 29–47.

Fodor, J. (2008). *LOT 2: The Language of Thought Revisited*. Oxford: Oxford University Press.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, *28*(1), 99–108.

Grice, P. H. (1975). *Studies in the way of words*. Cambridge: Harvard University Press.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, *87*(1).

Halberda, J. (2006). Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology*, *53*(4), 310–344.

Hill, A., Collier-Baker, E., & Suddendorf, T. (2012). Inferential reasoning by exclusion in children (Homo sapiens). *Journal of Comparative Psychology*, *126*(3), 243–254.

Horn, L. R. (1989). *A Natural History of Negation*. Chicago: University of Chicago Press.

Horst, J. S., & Samuelson, L. K. (2008). Fast Mapping But Poor Retention in 24-Month-Old Infants., (776112728).

Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy*, *15*(2), 125–150.

Itti, L., Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews. Neuroscience*, *2*(March), 194–203.

Jennings, R. E. (1994). *The Genealogy of Disjunction*. New York: Oxford University Press.

Laurence, S., & Margolis, E. (2002). Radical concept nativism. *Cognition*, *86*(1), 25–55.

Lea, R. B. (1995). On-line evidence for elaborative logical inferences in text. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *21*(6), 1469–1482.

Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: Developing "what" and "where" systems. *Trends in Cognitive Sciences*.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766.

Margolis, E., & Laurence, S. (2011). Learning Matters: The Role of Learning in Concept. *Mind and Language*, *26*(5), 507–539.

Markman, E., & Wachtel, G. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157.

Markman, E., Wasow, J., & Hansen, M. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, *47*(3), 241–275.

Mather, E., & Plunkett, K. (2011). Mutual exclusivity and phonological novelty constrain word learning at 16 months. *Journal of Child Language*, *38*(05), 933–950.

Mather, E., & Plunkett, K. (2012). The Role of Novelty in Early Word Learning. *Cognitive Science*, *36*(7), 1157–1177.

Merriman, W., Marazita, J., & Jarvis, L. (1995). Children's disposition to map new words onto new referents. In *Beyond names for things: Young children's acquisition of verbs* (pp. 147–183).

Mervis, C. B., & Bertrand, J. (1995). Acquisition of the novel name-nameless category (N3C) principle by young children who have Down syndrome. *American Journal of Mental Retardation: AJMR*, *100*(3), 231–243.

Munakata, Y. (2000). Challenges to the Violation-of-Expectation Paradigm: Throwing the Conceptual Baby Out With the Perceptual Processing Bathwater? *Infancy*, *1*(4), 471–477.

Nordmeyer, A. E., & Frank, M. C. (2014). The role of context in young children's comprehension of negation. *Journal of Memory and Language*, *77*, 25–39.

Peacocke, C. (1992). *A Study of Concepts*. Cambridge: MIT Press.

Peacocke, C. (2004). Interrelations: Concepts, Knowledge, Reference and Structure. *Mind and Language*.

Penn, D. C., & Povinelli, D. J. (2007). Causal cognition in human and nonhuman animals: a comparative, critical review. *Annual Review of Psychology*, *58*, 97–118.

Piaget, J. (1953). *Logic and Psychology*. Manchester: Manchester University Press.

Piaget, J. (1960). *The Psychology of Intelligence*. Totowa, NJ: Littlefield Adams & Co.

Premack, D., & Premack, a J. (1994). Levels of causal understanding in chimpanzees and children. *Cognition*, *50*(1-3), 347–362.

Quine, W. V. 0. (1960). *Word and object*. Cambridge: MIT Press.

Reverberi, C., Pischedda, D., Burigo, M., & Cherubini, P. (2012). Deduction without awareness. *Acta Psychologica*, *139*(1), 244–253.

Rey, G. (2014). Innate and Learned: Carey, Mad Dog Nativism, and the Poverty of Stimuli and Analogies (Yet Again). *Mind and Language*, *29*(2), 109–132.

Stalnaker, R. (1987). *Inquiry*. Chambridge: MIT Press.

Surian, L., & Caldi, S. (2010). Infants' individuation of agents and inert objects. *Developmental Science*, *13*(1), 143–150.

Téglás, E. (2010). From logic to probability: an investigation into the origins of probabilistic reasoning.

Tremoulet, P. D., Leslie, A. M., & Hall, D. G. (2000). Infant individuation and identification of objects. *Cognitive Development*, *15*(4), 499–522.

Watson, J. S., Gergely, G., Csanyi, V., Topal, J., Gacsi, M., & Sarkozi, Z. (2001). Distinguishing logic from association in the solution of an invisible displacement task by children (Homo sapiens) and dogs (Canis familiaris):

using negation of disjunction. *Journal of Comparative Psychology*, *115*(3), 219–226.

Xu, F. (2007). Sortal concepts, object individuation, and language. *Trends in Cognitive Sciences*, *11*(9), 400–406.

Xu, F., & Carey, S. (1996). Infants' metaphysics: the case of numerical identity. *Cognitive Psychology*, *30*(2), 111–153.

Xu, F., Cote, M., & Baker, A. (2005). Labeling guides object individuation in 12-month-old infants. *Psychological Science*, *16*(5), 372–377.

Zosh, J., & Feigenson, L. (2009). Beyond " what " and " how many ": Capacity , complexity , and resolution of infants ' object representations. *The Origins of Object Knowledge*, (c), 25–52.

Zosh, J. M., & Feigenson, L. (2012). Memory load affects object individuation in 18-month-old infants. *Journal of Experimental Child Psychology*, *113*(3), 322–336.

APPENDIX1: Familiarization movies

1. F1a

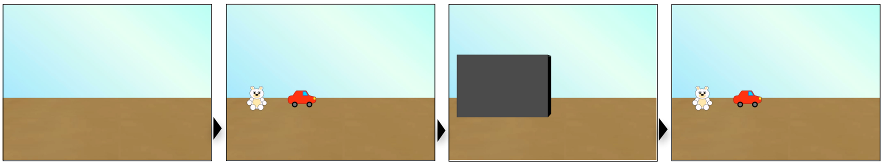

Figure 34. Schema of the familiarization movie F1a.

The movie begins with an empty scene. After 0.35 s, two objects (a white bear and a red car) enter the scene one after the other by falling from the upper edge of the screen. Both objects land in the left side of the scene (in describing the stimuli directions are always used relatively to the point of view of the viewer), on either side of each other. Immediately after the second object lands on the right of the first object, the two objects sequentially execute a characteristic movement. That is, the first object pulsates for 2 s while the second is motionless, and, as soon as the first object has stopped pulsating, the second object starts to pulse for 2 s while the first is motionless. The pulsating of each object is accompanied by a "cuckoo" sound. After the characteristic movements, a dark gray occluder emerges from the ground below the two objects and rises up to entirely cover the two objects. The two objects remain fully occluded for about 1 s. Afterwards, the occluder falls and disappears under the ground, revealing both objects in the same state as before the occlusion. After about 1 s, the two objects leave the scene, exiting sequentially through the left edge of the movie frame. The total length of the movie is 12.31 s.

2. F2a and F3a

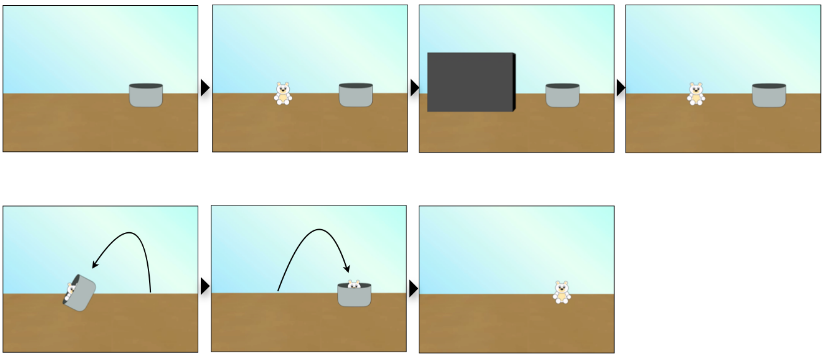

Figure 35. Schema of the familiarization movies F2a e F3a.

Each movie begins with an empty gray cup that lies in the right region of the scene. After 0.35 s, one object (a white bear in F2a and a red car in F3a) enters the scene by falling from the upper edge of the movie frame. The object lands in the left side of the scene. Immediately after the object lands, it executes its characteristic movement (i.e. the object pulsates for 2 s while the cuckoo sound is played). After the characteristic movement, a dark grey occluder emerges from the portion of ground below the object and rises up to entirely cover the object. The object remains fully occluded for about 1 s. Afterwards, the occluder falls and disappears under the ground, revealing the object in the same state as before the occlusion. Immediately after, the cup flies over the object and picks it up, bringing it to the right region of the screen. Now the cup is motionless and the upper part of the object is visible from inside the cup; where it again executes its characteristic movement. Afterwards, the cup reveals the object by falling and disappearing under the ground. The object, now completely visible, executes its characteristic movement. The total length of each movie is 15.81 s.

3. F4a

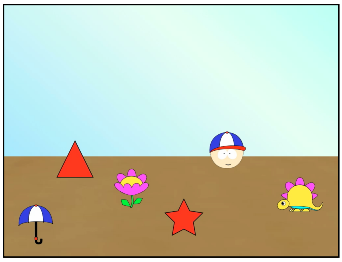

Figure 36. Screenshot of the familiarization movies F4a.

In the movie all six test objects (the objects that were used in the test movies) appear together in the scene. Starting from the object closest to the left edge of the movie frame, one after another, each object executes its characteristic movement. After the object closest to the right edge of the movie frame has completed its characteristic movement, the movie ends. The total length of the movie is 13.00 s.

4. F1b and F2b

The movies F1b and F2b are identical to the movies Familiarization F2a and F3a, except for the following changes. The movies introduce the functions of the cup and of the occluder when the two "act" separately (i.e. the object was taken by the cup only after the removal of the occluder). The white bear and the red car were replaced with two new objects: a red puppet with a head that culminated in a triangular tip and a red star with a multi-chromatic geometrical texture. (see Figure 37).

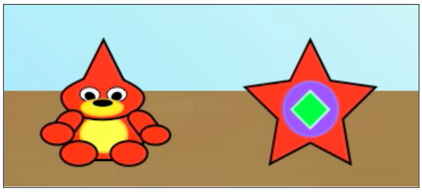

Figure 37. The pair Red-puppet/Star.

The two new familiarization objects had their characteristic movements. The red puppet moved its arms up and down several times. The red star rotated clockwise around its center. The red puppet and the red star could be used to produce the ambiguous containment: since they had an identical upper part; if partially hidden inside the cup, they look identical. Finally, certain events in the movies were slightly faster than in the movies F2a and F3a. The total length of each movie is 12.73 s.

5. F3b, F4b, F5b and F6b

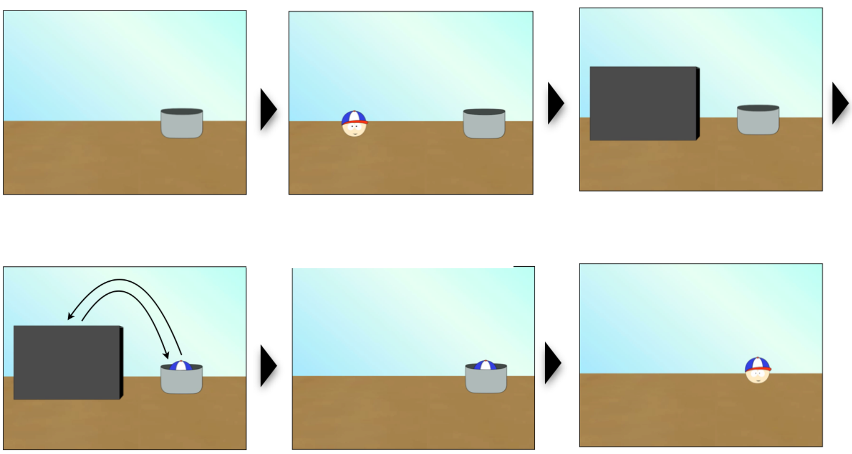

Figure 38. Schema of the familiarization movies F3b, F4b,F5b and F6b.

Each movie begins with an empty grey cup that lies in the right region of the frame. After 0.10 s, one of the test objects (the face in F3b, the umbrella in F4b, the dinosaur in F5b and the flower in F6b) enters the scene by falling from the upper edge of the movie frame. The object lands in the left side of the scene. Immediately after the object lands, it executes its characteristic movement accompanied by the cuckoo sound. After the characteristic movement, a dark grey occluder emerges from the portion of ground below the object and rises to cover the lower half part of the object. In this position the occluder leaves only the upper half of the object visible, the same part that was visible when the object was partially hidden inside the cup. The occluder remains stationary in this state for 1 s and afterwards it rises to entirely cover the object. As soon as the occluder has completely risen, the cup flies behind the occluder, entering from the center of its upper edge, and reemerges with the object inside, eventually landing in the right region of the scene. The cup remains motionless and the upper part of the object is visible from inside the cup. Afterwards, the occluder falls and disappears under the ground, revealing an empty region (i.e. no object remains behind the occluder). The region behind the occluder is revealed and, 0.30 s later, the object pulsates from inside the cup while the cuckoo sound is played. Immediately after the object inside the cup has completed the movement, the cup reveals its content by falling and disappearing under the ground. The object retrieved inside the cup is identical to the object initially covered by the occluder. The total length of each movie is 10.38 s.

6. F7b and F8b

Each movie had the same structure as the test movies in Experiment 3 (see section 3.2.1.2. Test Movies, for a description of the test movies), except for two features. First, in both movies we used the familiarization pair red puppet/red star (see APPENDIX1, section 4). Secondly, in both movies the outcome was missing. That is, after the object retrieved behind the occluder left the scene, the content of the cup was not revealed. Instead, the cup remained static on screen with the object hidden inside for 1 s and then the movie ends. In other words, the structure of the story presented in F7b and F8b is identical to that of a test movie except for the missing outcome: the content of the cup was not revealed. The movies F7b and F8b differed from each other only in which object was eventually retrieved from behind the occluder: in one movie it was the red puppet, in the other, the red star. The total length of each movie is 17.98 s.

7. F1c and F2c

The two movies are identical to F1b and F2b (see APPENDIX1, section 4), respectively, except for the timing of certain events. The total length of each movie is 13.48 s.

8. F3c and F4c

The two movies have the same structure as F3b, F4b, F5b and F6b (see APPENDIX1, section 5), except for the timing of certain events and the use of the red puppet and the red star rather than the test objects. The total length of each movie is 11.73 s.

9. F5c and F6c

The two movies are identical to F7b and F8b (see APPENDIX1, section 6), respectively, except for the timing of certain events. The total length of each movie is 18.98 s.

10. F1d and F2d

In the movies we used a new object pair: Blue-puppet/Blue-flower (See Figure 39).

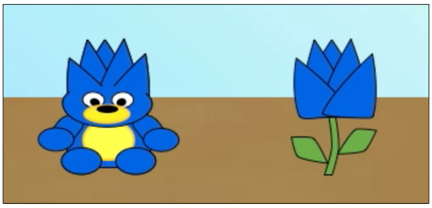

Figure 39. The pair Blue-puppet/Blue-flower.

The two new familiarization objects have their characteristic movements: the blue puppet moves its arms up and down several times, while the blue flower oscillates left and right several times. The blue puppet and the blue flower pair were designed to be used to produce the ambiguous containment; they have an identical upper part so that, if partially hidden inside the cup, they look identical.

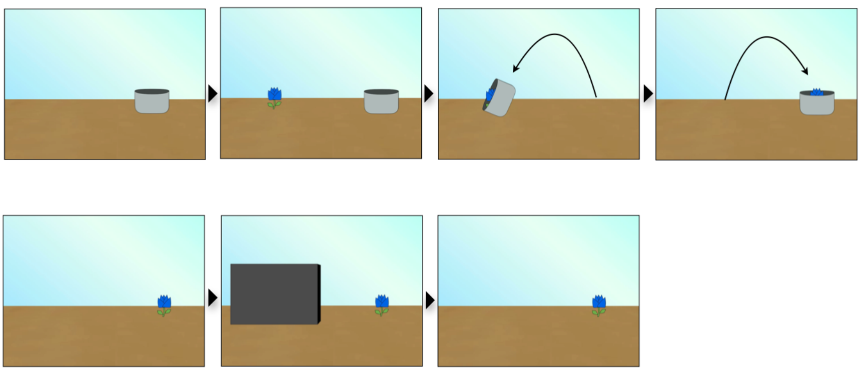

Figure 40. Schema of the familiarization movies F1d and F2d.

Each movie begins with an empty gray cup that lies in the right region of the scene. After 0.20 s, one object (the blue puppet in F1d and the blue flower in F2d) enters the scene by falling from the upper edge of the movie frame. The object lands in the left side of the scene. Immediately after the object lands, it executes its characteristic movement. After the characteristic movement, the cup flies over the object and picks it up, bringing it to the right region of the screen. The cup then reveals the object by falling and disappearing under the ground. The object, now completely visible, executes its characteristic movement. Immediately after, a dark grey occluder rises from the portion of ground below the empty left region of the scene. The occluder remains fully risen for about 1 s. Afterwards, the occluder falls and disappears under the ground, revealing the empty left region of the scene. The total length of each movie is 13.48 s.

11. F3d and F4d

The movies were identical to F3c and F4c, except for two changes. First, the objects used were the blue puppet (in F3d) and the blue flower (in F4d). Secondly, the occluder is not removed. The total length of each movie is 15.36 s.

12. F5d and F6d

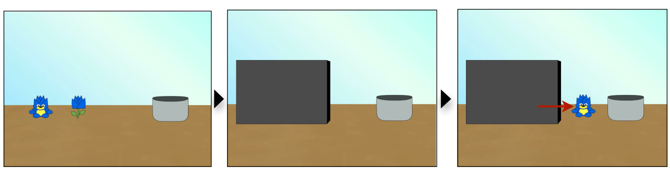

Figure 41. Schema of the familiarization movies F5d and F6d.

The movies begin with an empty cup that lies in the right region of the scene. After 0.20 s, two objects enter the scene, one after the other, by falling from the upper side of the movie frame. Both objects land in the left side of the scene, on either side of each other. Immediately after the second object has landed on the right side of the first object, the two objects sequentially execute their characteristic movements. After the objects' characteristic movements, the occluder rises up to cover the lower half of the objects, leaving visible only their upper half. The occluder remains stationary in this state for 1 s and then the occluder falls and disappears under the ground, revealing both objects in the same state as before the occlusion. The objects remain entirely visible for about 1 s and then the occluder rises again to cover the lower half of the objects. The occluder remains stationary in this state for 1 s and then rises to entirely cover the two objects. Afterwards, one of the two objects (the blue puppet in F5d, the blue flower in F6d) exits from the right edge of the occluder, stops between the occluder and the cup, and executes its characteristic movement. The total length of each movie is 14.79 s.

13. F1e and F2e

The two movies are identical to F1b and F2b (see APPENDIX1, section 4), respectively, except for certain events that were slightly faster. The total length of each movie is 13.48 s.

14. F3e and F6e

The two movies are identical to F3b and F6b (see APPENDIX1, section 4), respectively, except that the occluder is not removed. The total length of each movie is 13.86 s.

15. F7e and F8e

Each movie has the same structure as the test movies of Experiment 6 (see the section 7.2.1.2. test movies, for a description of the test movies) except for two aspects. First, in both movies we used the familiarization pair red puppet/red star (see APPENDIX1, section 4). Secondly, in both movies the outcome was missing. That is, after the first object has exited from behind the occluder and returned behind it, no other object exits from behind the occluder. Instead, the cup remains static on the screen with the object hidden inside for 1 s and then the movie ends. In other words, the structure of the story presented in F7e and F8e was identical to that of the test movies except for the missing outcome: one object exits from behind the occluder only once. The movies F7e and F8e differed from each other only in which object exited from behind the occluder: in one movie it was the red puppet, in the other the red star. The total length of each movie is 17.98 s.

APPENDIX2: Preliminary Analyses of Distance and Pair Effects

For each experiment from Experiment 1 to 6, before collapsing looking time trial scores by averaging them by Outcome Type (Consistent vs Inconsistent), we performed preliminary analyses to assess the effects of the object pair used in the trial, and the initial distance from the cup to the object eventually retrieved inside the cup. The following table (see Table 6) summarize the outcomes of the analyses.

Experiment	Effect	F-Ratio	Probability
Experiment 1-A	Distance	$F(1,40) = 0.04$	$P = 0.83$
Experiment 1-A	OutcomeType	$F(1,40) = 6.94$	$P < 0.05$
Experiment 1-A	Distance*OutcomeType	$F(1,40) = 0.18$	$P = 0.67$
Experiment 1-B	Pair	$F(2, 81) = 2.63$	$P = 0.06$
Experiment 1-B	OutcomeType	(2, 81) = 9.16	$P < 0.005$
Experiment 1-B	Pair*OutcomeType	$F(2, 81) = 0.29$	$P = 0.75$
Experiment 2-A	Distance	$F(1, 44) = 0.00$	$P = 0.95$
Experiment 2-A	OutcomeType	$F(1, 44) = 0.21$	$P = 0.65$
Experiment 2-A	Distance*OutcomeType	$F(1, 44) = 3.96$	$P = 0.05$
Experiment 2-B	Pair	$F(2, 93) = 2.63$	$P = 0.08$
Experiment 2-B	OutcomeType	$F(2, 93) = 0.78$	$P = 0.38$
Experiment 2-B	Pair*OutcomeType	$F(2, 93) = 0.3$	$P = 0.74$
Experiment 3-A	Distance	$F(1,42) = 1.08$	$P = 0.3$
Experiment 3-A	OutcomeType	$F(1,42) = 5.43$	$P < 0.05$
Experiment 3-A	Distance*OutcomeType	$F(1, 42) = 5.64$	$P < 0.05$
Experiment 3-B	Pair	$F(1, 60) = 7.66$	$P < 0.01$
Experiment 3-B	OutcomeType	$F(1, 60) = 8.32$	$P < 0.01$
Experiment 3-B	Pair*OutcomeType	$F(1, 60) = 4.43$	$P < 0.05$

Experiment	Effect	F-Ratio	Probability
Experiment 4-A	Distance	$F(1,42) = 0.83$	$P = 0.36$
Experiment 4-A	OutcomeType	$F(1,42) = 2.05$	$P = 0.16$
Experiment 4-A	Distance*OutcomeType	$F(1, 42) = 0.00$	$P = 0.95$
Experiment 4-B	Pair	$F(1, 56) = 0.14$	$P = 0.7$
Experiment 4-B	OutcomeType	$F(1, 56) = 1.78$	$P = 0.19$
Experiment 4-B	Pair*OutcomeType	$F(1, 56) = 1.14$	$P = 0.29$
Experiment 5-A	Distance	$F(1,44) = 0.24$	$P = 0.62$
Experiment 5-A	OutcomeType	$F(1,44) = 3.37$	$P = 0.073$
Experiment 5-A	Distance*OutcomeType	$F(1, 44) = 0.21$	$P = 0.65$
Experiment 5-B	Pair	$F(2, 88) = 0.39$	$P = 0.68$
Experiment 5-B	OutcomeType	$F(2, 88) = 5.16$	$P = 0.026$
Experiment 5-B	Pair*OutcomeType	$F(2, 88) = 2.3$	$P = 0.11$
Experiment 6-A	Distance	$F(1,44) = 0.00$	$P = 0.95$
Experiment 6-A	OutcomeType	$F(1,44) = 7.62$	$P < 0.01$
Experiment 6-A	Distance*OutcomeType	$F(1, 44) = 1.35$	$P = 0.25$
Experiment 6-B	Pair	$F(1, 46) = 0.91$	$P = 0.34$
Experiment 6-B	OutcomeType	$F(1, 46) = 7.76$	$P < 0.01$
Experiment 6-B	Pair*OutcomeType	$F(1, 46) = 0.62$	$P = 0.43$

Table 6. Summary of the results of the preliminary analyses of Experiments 1 to 6. For each experiment, two separate ANOVAs were performed. One two-way ANOVA (marked with "A" in the summary table) with Outcome Type and Distance (Close vs Distant) as between participant factors and Participant as a random variable nested in both Outcome and Distance was performed to assess the possible effects of Distance. One two-way ANOVA (marked with "B" in the summary table) with Outcome Type and Pair as a within-participants factors and Participant as random factor was performed to assess the possible effects of Pair.