This is how we do it: Interpreting sub-efficient means actions of group members in infancy and childhood

Nazlı Altınok

Submitted to:

Central European University

Department of Cognitive Science

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Cognitive Science

Primary supervisor: György Gergely

Secondary supervisor: Ildikó Király

External advisor: Mikołaj Hernik

Budapest, Hungary

2020

Declaration of Authorship

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or which have been accepted for the award of any other degree or diploma at Central European University or any other educational institution, except where due acknowledgment is made in the form of bibliographical reference.

The present thesis includes work that has appeared or will appear in the following articles:

Altınok, N., Hernik, M., Király, I., & Gergely, G. (2020). Acquiring sub-efficient and efficient variants of novel means by integrating information from multiple social models in preschoolers. *Journal of Experimental Child Psychology*, *195*, 104847.

Altınok, N., Király, I., & Gergely, G. (in preparation). Eighteen-month-olds' selective imitation of demonstrators who speak the same language as the infant's social community: Efficiency does not always triumph.

Altınok, N., Hernik, M., Király, I., & Gergely, G. (in preparation). Eleven-month-olds expect agents to act efficiently but also to repeat the movements of their group members.

Nazli Altinok

Nazlı Altınok

Abstract

Human children must acquire culture-specific conventional practices of their social groups to be competent members of their own communities. Given that such conventional knowledge forms are often cognitively opaque to be learned through individual exploration alone, in order to acquire cultural knowledge of their communities, young children must largely rely on the evidence provided by knowledgeable others and be ready to faithfully imitate seemingly inefficient and arbitrary means actions. The present work addresses whether representing social groups facilitates how infants and young children interpret, acquire and maintain sub-efficient means actions.

The first study investigated whether 18-month-old infants selectively imitate cognitively opaque actions from models who speak the same language as their own language community in ostensive demonstration contexts. The main finding was that infants showed selective imitation of sub-efficient action demonstrations as a function of the language the demonstrators spoke. The second study explored whether 14- and 18-month old infants expect a particular opaque action form to be shared only among the members of the same social group. Neither 14-month-olds nor 18-month-olds were found to form expectations about how other group-members ought to act on a novel apparatus based on the prior ostensive demonstration of the sub-efficient action manner by a model who belonged to the same linguistic group as the infants and as the other agents in the scene. The third study probed pre-verbal infants' expectation of shared movement repertoire between animated characters belonging to the same social group and how this shared-movement expectation interacts with their efficiency expectation. The findings showed that expectation of a shared movement repertoire between agents acting in social groups influences 11-month-old infants' expectations of efficiency of goal directed actions. The final study explored whether preschoolers, in three age groups (4-, 5- and 6-year-olds), can treat sub-efficient action forms and their efficient alternatives in a

context-sensitive selective manner, depending on the social group membership of the demonstrator. Four-year-olds employed the cue of shared language to optimize acquiring and maintaining culturally shared sub-efficient action routines by selectively updating their action repertoire relying on their language-based evaluation of the demonstrator's culture-specific competence. In contrast, 5- and 6-year-olds adopted the efficient alternative independently of the demonstrator's language.

The work presented in this thesis has built on different methodologies across different age groups to investigate how sub-efficient action routines are represented, acquired or maintained. Beginning from the age of 11-months infants show sensitivity to the shared-movement repertoires of social groups. Before their second birthday they readily acquire sub-efficient routines from the informants speaking the same language as them. Around the age of 4 years they also rely on the informants' language as a cue of shared cultural ground and incorporate efficient alternatives into their knowledge repertoire depending on whether the informant speaks the same language as them. Taken together, this work extends previous research on children's social learning and early action understanding and provides evidence for how we use social categories from an early age on in order to make sense of the behaviors of others.

To the beloved memory of my father,

Acknowledgments

I would first like to thank to György Gergely for his valuable supervision and his theoretical guidance during this endeavor. Similarly, I would like to thank to Ildikó Király and Mikołaj Hernik for their unwavering support and immensely constructive mentorship through the years. My words fall short for the gratitude I have for you all. Ildikó, without your encouragement this would not have been possible. Thank you.

I would also like to thank to all CDC members, and all the amazing research assistants I have been privileged to know: Iulia Savoş, Dóri Mészégető, Zsuzsanna Karap, Szilvia Takács, and Petra Kármán. Thank you to our coordinators Maria Toth and Agi Volein for helping me to maneuver my way around the babies in a language that I am still learning. *Köszönöm* to Réka Finta, Andrea Schrök, Kati Illés, and Edit Vitrai for their support in the departmental life. I am indebted to Gergő Csibra and Ági Volein, for having an open door for me whenever I needed and for all the support they have kind-heartedly provided. To Ági Kovács, thank you for reading my work, for your insightful comments and for your encouragement. To Malinda Carpenter, thank you for hosting my research visit to the University of St Andrews, and helping me to grow as a person and a researcher. Paula Rubio-Fernández: your support meant a lot to me in every aspect of my academic life, thank you. I would also like to thank to Christophe Heintz, Natalie Sebanz, Olivier Mascaro and Ernő Téglás for their guidance.

To all my friends and colleagues in the department, especially Luke, Laura, Mia, Oana, Otavio, Lenka, Denis, Hanna, Helena, Tibor, Eszter S., Ieva, Arianna, Martin S., Martin D., and Çağlan: thank you for your support and encouragement over the years. I am lucky to know each one of you. Thank you to my friends: Ayşegül, Baran, Ceren, Ceylan, Defne, Eymen and Mustafa. Not only for opening your doors to me during my conference trips from Boston to Glasgow but to make the time and effort to me be on my side whenever I needed it, I am grateful for your friendship.

I would also like to thank the hundreds of families participating in this research. This work would not be possible without their interest in developmental science.

Lastly, I would like to thank my mother, whose resilience and hard work motivates me to better myself every day. You have been my cheerleader through my academic and personal growth. *Teşekkürler annem*.

I dedicate this thesis to the memory of my beloved father, you fought well and taught me a lot about life, death, and everything in between. It pains me not being able to share all this with you, but I know that you would be proud.

Table of contents

Chapter I. General Introduction

In Kachin customary procedure the routines of clearing the ground, planting the seed, fencing the plot and weeding the growing crop are all patterned according to formal convention and interspersed with all kinds of technologically superfluous frills and decorations which make the performance a Kachin performance and not just a simple functional act. And so it is with every kind of technical action; there is always the element which is functionally essential, and another element which is simply the local custom, an aesthetic frill.

Leach, 1954, p. 12

Human children must acquire a large body of cumulative cultural knowledge in a variety of domains from learning about artifact functions to acquiring language, social conventions and normative behaviors of their cultural community. Given that such knowledge forms are often too complex to be learned through individual exploration alone (Gergely & Csibra, 2006; Legare & Nielsen, 2015; Legare, 2019), in order to acquire and preserve cultural knowledge, infants and young children must largely rely on the evidence provided by the observed intentional actions and/or communicative demonstrations of knowledgeable others. Social learning is a central mechanism that enables novices to learn new skills and knowledge (Boyd, Richerson, & Henrich, 2011; Tomasello, 1996, 1999a; Gergely & Csibra, 2006). However, not all forms of social learning could enable children to be efficient cultural learners. For example, in emulative learning (Tomasello, 1996), an individual could observe an actor bringing about an outcome and can reliably reproduce the same outcome herself later by using a different strategy. Emulation (i.e. copying the results of the observed actions, Call & Carpenter, 2002) is one of the key social learning mechanisms enabling novices to discover the affordances of the environment that they might not be likely to discover by themselves. However, emulative learning by itself would fall short ensuring cultural transmission and maintenance of certain cultural knowledge forms that are cognitively opaque (Gergely &

Csibra, 2006). For instance, in learning of conventional practices, the naïve learner cannot rely on emulation as there is often no underlying physical causal mechanism (neither an apparent adherence to principles of efficiency of means) that would make the functions of such conventional actions obvious to the observer. To acquire such cognitively "opaque" (sub-efficient) cultural procedures naïve leaners need to closely adhere to the exact way these actions are performed by the competent experts in the cultural community and engage in faithful imitation.

Various cognitive biases support high-fidelity imitation. Csibra & Gergely (2009, 2011), in their theory of Natural Pedagogy, argued that our uniquely human propensity to attend to ostensive communicative signals (such as eye contact, infant directed speech, contingent reactivity), indicating the teaching intention of others, enables us to construe the content of the information as relevant, novel, and generic, shared by the social community. The authors also argued that our evolved propensity to differentially attend to ostensively demonstrated action sequences is fitness enhancing as it promotes the transfer of complex cultural knowledge forms with opaque causal or teleological properties that the naïve learner would be unable or unlikely to discover through trial-and-error learning on her own. In fact, many argue (Boyds & Richerdson, 1985; Sperber, 1996; Tennie, Call, & Tomasello, 2009; Butler & Markman, 2012; Nielsen, 2012) that human's species-unique ability to inter-generationally transmit and faithfully reproduce complex cumulative cultural knowledge has been made possible by the selection during hominid evolution of dedicated cognitive adaptations involving specialized learning mechanisms that enable the faithful transmission of a wide range of novel and complex cultural skills even when their content appear causally and functionally opaque to the naïve learner (Csibra & Gergely, 2009, 2011, Gergely & Csibra, 2005, 2006).

However, cultural knowledge is bounded by social groups. In social groups members are likely to have common history together, and consequently they might have shared preferences, attitudes, and knowledge forms that distinguishes them from other social groups. As naïve learners socializing into a certain social group, infants and children should be epistemically attuned to their own cultural community. They should not simply acquire and persevere with every instance of ostensively communicated cultural forms of knowledge but should trust and rely on the members of their community to ensure that they are learning the cultural knowledge repertoire relevant for their own communities. While doing so they also should be sensitive to the boundaries of cultural knowledge and should not expect cultural knowledge of their own community to be available to those from another community (see Diesendruck & Markson, 2011; Diesendruck, Carmel, & Markson, 2010, on conventionality). Based on these premises, I hypothesize that representing social groups facilitates the transmission and maintenance of some cultural knowledge forms, such as conventional practices, which are part of the body of cumulative cultural knowledge that are inter-generationally transmitted in human social groups. Furthermore, such cultural knowledge forms could serve a social function of ingroup identification by providing salient "group markers", acting as reliable signals displaying one's commitment to the practices of the group. In this reading, shared cultural repertoire between agents could be a cue revealing that they belong to the same social group, guiding naïve learners' predictions and inferences about what kind of actions agents from the same group *ought to* perform in certain social contexts.

Considering the above hypotheses, the empirical chapters in this thesis explore the following main questions:

1) Do infants show selective imitation of cognitively opaque actions (i.e. "actions that are unpredictable on the grounds of physical-causal efficiency", Gergely & Csibra, 2005, p. 470) if such actions are presented in ostensive demonstration contexts by

models who belong to the same social group as the infants themselves? (testing for transmission of cultural knowledge, Chapter 2)

2) If so, would they expect a particular cognitively opaque action form to be shared only among the members of the same social group? (testing for the sharedness and generalizability of cultural knowledge, Chapters 3 & 4)
3) How do older children retain cognitively opaque actions despite observing their efficient alternatives? (testing for the maintenance of cultural knowledge, Chapter 5)

The present introductory chapter will provide a review of the available literature and empirical results that are relevant for the questions articulated above, while later chapters will report how I have addressed each of them experimentally. This chapter will first cover previous studies investigating how infants make sense of the actions they observe. Note that most of these studies, if not all, focused on instrumental actions only. Then I will introduce the theoretical concept of a "ritual stance" by outlining how socially stipulated, teleologically or causally opaque actions are discussed in the literature by citing studies carried out with adults, and with young children. After inviting the reader to consider the challenges that naïve learners might potentially face interpreting, acquiring, and preserving such cognitively opaque actions, which I consider as a central form of shared "cultural knowledge", I will review the relevant literature on social cues infants and young children can depend on while solving this learning challenge. I will introduce social group membership cues and their role in signaling the informants' possession of relevant cultural knowledge. My proposal is that naïve social learners should exploit and rely on cues that are indicative of social group membership while inferring other parties' cultural knowledge repertoires in observational settings, and also while acquiring and preserving cultural knowledge. After all, relying on someone's communicative intention alone might not guarantee that the informative content of the communicative intention is the most relevant piece of knowledge for a novice to acquire in every context. Therefore, naïve learners should be able to exploit more direct cues of relevance if available, when it comes to acquiring cultural knowledge from others. Cues indicative of being a member of the same social group as the naïve learner is an informative indicator ensuring that the presented new information is, indeed, relevant (and shared) cultural knowledge. Hence in the last sections of this chapter, I will introduce the hypothesis that cues indicating same group membership also function as informative signals about the informant's possession of relevant cultural knowledge. The final section of this introductory chapter will, therefore, review relevant studies from the literature on group membership and discuss the functions of social categories in learning.

The work presented in this thesis aims to build on, but also to extend previous research on children's social learning, studies related to early action understanding, and experimental work on social group cognition. To this end, Chapter 2 will investigate 18-month-old infants' selective imitation of models speaking the infant's own language in acquiring cognitively opaque, sub-efficient means-actions (henceforth "cultural knowledge"). Chapter 3 will present experimental work on 14- and 18-month-old infants' attribution of shared cultural knowledge to the speakers of their language versus a foreign language in a looking time paradigm. Chapter 4 will report experimental studies on 11-month-old infants' expectation of shared movement repertoires among the same group members, in another looking time paradigm. Finally, applying a modified version of so-called "over-imitation" procedures (Horner & Whiten, 2005; Lyons, Damrosch, Lin, Macris, & Keil, 2011; Lyons, Young, & Keil, 2007) I will present studies to explore how older children, preschoolers, in three age groups (4-, 5- and 6-year-olds) acquire and retain cultural knowledge as a function of the demonstrator's social group membership in Chapter 5. In the next section I will provide an overview of the current state of the art in developmental research on action understanding and social group cognition, serving as a backdrop for the research presented in subsequent chapters.

1.1 How do we interpret actions: the teleological stance

Human infants are endowed with an interpretation system that allows them to make predictions and inferences about actions using the principle of rationality: given the environmental constraints, agents are expected to obtain their goals through the most efficient action means available (the teleological stance, or the naïve theory of rational action, see Gergely & Csibra, 2003; Csibra & Gergely, 2007). Numerous violation-of-expectation looking time studies corroborated the hypothesis that infants from early on interpret actions as goal-directed by applying the teleological stance. In a violation of expectation paradigm, 12-month-olds were habituated to an event where a ball jumped over a long barrier to reach another small ball on the other side of the screen. At test, the barrier was removed allowing the ball to more directly access the small ball following a straight trajectory and infants viewed two events: one test event depicted the ball pursuing a direct path towards its goal (visually novel scene in comparison to what infants were habituated to, but given the removal of the barrier it is a more efficient action), while the other test event depicted the ball repeating the exact same jumping movement the infants were habituated to despite the barrier now not being present (visually similar scene to the habituation, but given the removal of the barrier it is an inefficient action). Findings revealed that infants looked longer (detecting a violation in their expectations of efficient goal-approach formed during habituation) to inefficient test trials in comparison with efficient test trials (Gergely, Nadasdy, Csibra, & Biro, 1995). This expectation of the most efficient path in goal-directed actions has been replicated in several other studies (Csibra, Gergely, Bíró, Koós, & Brockband, 1999; Biro, Verschoor, & Coenen, 2011, Csibra, 2008a) in a multitude of different contexts, for example not only using animated agents, but also puppets (Sodian, Schoeppner, & Metz, 2004), hands (Philips & Wellmann, 2005), robots (Kamewari, Kato, Kanda, Ishiguro, & Hiraki, 2005), different action trajectories (Liu & Spelke, 2017), and with infants as young as 3 month of age (Liu, Brooks, & Spelke, 2019). The teleological interpretation system has also been demonstrated in in some non-human primate species (Rochat, Serra, Fadiga, & Gallese, 2008; Uller, 2004).

Empirical evidence reliably documented that infants do not have to have the action in their own motor repertoires to make sense of it, rather the teleological interpretation mechanism they are endowed with allows them to make sense of others' actions even when the action is novel or the agent performing it is unfamiliar[1]. The teleological interpretation system succeeds in doing so as the link between the agent, the action and the physical context is governed by the inferential principle of rationality (or efficiency). Infants can predict the action an agent is going to take given her or his goal and the physical constraints of the environment. Further they can also infer physical goals from observing the action alone since they assume that the action would be the most optimal way of achieving the goal that they did not yet observe, given the environmental constraints (Brandone & Wellmann, 2009; Csibra et al. 2003; Daum, Prinz, & Aschersleben, 2008; Southgate & Csibra, 2009; Wagner & Carey, 2005; Meltzoff, 1995). Additionally, the teleological inferential system allows infants to infer the unseen physical constraints in the environment to make sense of why an agent apparently deviated from the most efficient action path to achieve her goal (i.e. by inferring the presence of an unseen obstacle, they can interpret the observed action as being a rational means to the goal). A direct

[1] Sommerville & Woodward (2005a) along with several others (Skerry, Carey, & Spelke, 2013) showed that prior experience facilitates goal encoding among pre-verbal infants. Also see mirroring accounts on action understanding (Rizzolatti & Craighero, 2004; Rizzolatti, Fogassi, & Gallese, 2001; Buccino, Binkofski, & Riggio, 2004). Here I do not argue that motor simulation and action familiarity have no role in action understanding. Prior experience can indeed facilitate action understanding as already documented, but it is not required for early action understanding and do not constitute the basis for infants' teleological stance in interpreting actions.

test for the latter inference comes from Csibra et al. (2003). Infants were habituated to events similar to those described in Gergely et al.'s (1995) task, with the exception that this time the middle section of the path the agent took was occluded during the habituation. During the alternating test trials, the occluder in the middle of the path was raised revealing either an empty space between the agent and its end-state (i.e. no-obstacle event) or a cube blocking the agent's straight path towards its end-state (i.e. obstacle event). When 12-month-old infants viewed the exact same jumping action which they were previously habituated to in these two test trials, they looked longer to the no-obstacle event than the obstacle event. This finding is a clear demonstration that infants can infer the presence of the unseen physical constraint in the environment to justify the agent's action (as efficient) towards its goal.

Productive inferences infants make based on the principle of rationality in this tripartite model of action interpretation are not limited to studies that make use of scenarios where infants are shown odd actions that deviate from learned statistical regularities in daily life (i.e. jumping over non-existing obstacles, for statistical learning account of action understanding see Paulus, Hunnius, van Wijngaarden, Vrins, van Rooij, & Bekkering 2011; Paulus, Hunnius, Vissers, & Bekkering, 2011a). Nor are they limited to studies that might simply tap into infants' expectation that the agents should take the shortest path possible while obtaining their goals, without actually having a general expectation of rationality. A study that made use of scenarios where an agent can either perform fewer actions to obtain a target item from a container with no cover (i.e. less effort) or she can perform more actions to obtain an identical looking item from another container with a transparent cover (pull the cover, take the toy) (i.e. more effort) also documented infants' surprise upon seeing the agent expending more effort in retrieving the object when the identical looking object could be more easily accessed as it was within her direct reach (Scott & Baillargeon, 2013). Another study also demonstrated that infants' expectation of rationality is not a byproduct of their surprise reaction to odd and infrequent events (Southgate, Johnson, & Csibra, 2008). Infants were familiarized to a human hand moving a box away in order to reach out and retrieve a ball that was obscured behind the box. In alternating test trials, infants saw the hand moving two boxes out of the way to retrieve the ball ("possible less efficient") or they saw a more efficient, but biomechanically impossible movement (i.e. going ahead in between the boxes in a "snake-like" manner – a biomechanically impossible movement for a human hand to perform) to retrieve the ball. Six- to eight-month-old infants looked longer at "possible, but less efficient" test trials indicating that they have interpreted the observed actions based on the principle of efficiency, even the when the rational action they saw was not only unfamiliar, but also biomechanically impossible.

Further evidence pointing to an expectation of efficiency comes from a selective imitation study (Gergely, Bekkering, & Kiraly, 2002) where the researchers slightly modified the original procedure of Meltzoff's (1988) head-touch study. In their version of the head-touch task, the demonstrator had wrapped a blanket around her shoulders causing her hands to be occupied before leaning on the touch-sensitive light box to get her forehead in touch with the lamp thereby bringing about a light effect. Upon being demonstrated the hands-occupied head-action the number of 14-month-old infants reenacting the head-action in the imitation phase dropped significantly in comparison to a condition where there was no such constraint to the demonstrator's action (hands-free head action), just like in Meltzoff's original study. In the hands-occupied condition, the infants' own hands were not occupied (and so they did not share the same constraints as the model), therefore, they did not imitate the head-action of the model with fidelity, because touching the box with head was not the most rational means available for them (as they could and did use their hands to operate the touch-lamp instead, Gergely et al., 2002; Kiraly, Csibra, & Gergely, 2013).

As summarized above several studies show that infants rely on the assumption of rationality while making sense of actions they observe from very early on. Without the assumption of rationality of goal-approach, infants would not be able to assign the goal, and understand and predict the goal-directed actions of other agents observed in their environment. However, the rationality of an action is not always constrained by the physical environment alone but also by the agents' mental states (e.g. preferences, desires) and the social context they are in (e.g. group norms and conventions). While the teleological stance identifies a powerful interpretive and computational mechanism by which infants can represent actions in terms of efficiency of physical goal-approach, it does not extend in a straightforward manner to explain how the perceived value of the end state and social contexts can guide an agent's action.

1.1.1 Interpreting actions beyond physical efficiency: the role of attributed dispositional properties of agents

An action interpretation system that is guided by the rationality assumption reliably explains an agent's behavior of approaching a goal object or bringing about a physical change of state in the world by breaking down the observed behavioral event into three elements: action, environmental constraints and goal (i.e. assuming that the agent's goal is to get to a goal-object she is expected to take the route that is the optimal, the most efficient [least costly] available within the constraints of her physical environment). But what if there are multiple other goal locations? The teleological stance might predict that the agent will approach the goal object that is located closer to her with the assumption that the other goal objects at the alternative locations are of the same value to the agent. However, one can imagine a myriad of scenarios, for example, where the far-away source contains a kind of goal object that the agent has a higher preference for, or where the closest goal source contains a potential goal object, which is however poisonous. End-states are not always neutral, and agents' preferences and knowledge states can guide their actions to obtain different goals as a function of their different attributed value selectively. As an observer I should be able to represent that the agent has a preference for the first type of goal object, and that the agent knows that the nearby source contains a potential goal object, that – due to its negative value – should be avoided. Is there evidence that early action understanding is also guided by these non-visible dispositional and value-related properties of the goal objects?

Scott and Baillargeon (2013) showed that 16-month-old infants expect an agent to retrieve the item she has the preference for even if it means expending more effort for obtaining it in contrast with the other item on the scene that is rendered more accessible to the agent (and could be retrieved by expending less effort) (Experiment 1, different-objects condition). When agents have a preference for a particular item, the degree of effort exerted to retrieve it can be justified by the infant, even when the action means required to obtain the goal involves more invested effort (it is more costly to perform). Such a more costly action means can be considered rational as it can be justified by the larger benefit obtained (due to the agent's subjective preference or her knowledge of the higher objective value of the goal object).

Recent studies showing that infants and young children could take agents' preferences into account in predicting future actions and identifying goal-states comes from research motivated by a computational version of the teleological account, called a "naïve utility calculus" (Jara-Ettinger, Gweon, Tenenbaum, & Schultz, 2015; Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016). A naïve utility calculus provides a framework incorporating the notion of utility into the schema of action planning by making the following assumptions: 1) agents aim to maximize utility; 2) utility is calculated based on expected rewards and costs, so that agents maximize anticipated rewards relative to expected costs; 3) the notion of cost depends on the properties of the agent and the situation: together they determine how much of an effort an agent needs to expend to perform the action (Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016; Jara-Ettinger, Floyd, Tenanbaum, & Schulz, 2017; Liu & Spelke, 2017). Several studies have documented that we expect others to be utility maximizers from early on, revealing that even at a young age we can consider the hidden properties of goals (e.g. value) and actions (e.g. physical and psychological effort) while interpreting others' behaviors.

The naïve utility calculus account predicts that agents will pursue an action plan only if expected rewards outweigh the expected costs. For example, in one study 5-year-old children were shown a puppet, who first chose crackers over cookies when both items were placed equally distant from the puppet. Later children saw the same puppet choosing cookies over crackers when cookies were placed closer. Upon being asked what item the puppet has a preference for 5-year-olds inferred the puppet's preference based on the choice she made in the initial scene, when the costs for the items were the same. In other words, they did not regard the puppet's choice as the same as her preference without considering the respective costs that achieving the respective goal items incurred (Jara-Ettinger et al., 2015). In a follow-up study, children were this time shown two puppets, one that liked crackers over cookies and the other with no preference. When the crackers were placed on a tall box, and when both puppets chose the cookies that were on the low box, the children were asked which puppet cannot climb. Five-year-olds inferred that the puppet who liked the crackers was the one that could not climb (Jara-Ettinger et al., 2015).

Naïve utility calculus provides a rich inferential system and broadens the non-mentalistic teleological action interpretation system by explaining how we think other people pursue their goals and how not-always-obvious and sometimes subjective properties (such as preference, value, competence) can guide early action understanding. Even pre-verbal infants were found to be sensitive to subjective and objective features of costs and rewards, expecting agents to maximize their utilities. In a violation of expectation paradigm, Liu and Spelke (2017) familiarized 10-month-old infants to scenes where an animated agent incurred low cost to approach a target (i.e. jumped over a low barrier to get closer to the target) but rejected to expend medium effort to approach the same target (i.e. did not jump over a medium-height barrier to get closer to the object while making a negative sound as she backed away from the barrier). During the familiarization, infants also watched the same agent incurring a medium cost to approach another target, but rejecting to expend high effort in her approach (i.e. the agent jumped over a medium-height barrier to get closer to another target but rejected to jump over a higher barrier). Following these familiarization events, infants expected the agent to approach the high-value target (the target the agent incurred the medium cost to reach) in contrast with the low-value target (the target the agent incurred only low cost to reach) when there were no obstacles in the scene. This finding constitutes clear evidence showing that pre-verbal infants can interpret the value of the target the agent approaches from the degree of effort she had expended to approach it earlier. A recent study also documented that 13-month-old infants can use the degree of peril in an action to infer the value of the goal for the agent performing that action. This study documents that infants also calculate and base their inferences on negative utilities in understanding others' actions, extending on earlier findings which demonstrated that infants can make use of different degrees of effort an agent invests to infer the value of the goal objects for that agent (Liu, Ullmann, Tenanbaum, & Spelke, 2019).

Under the assumptions of the naïve utility calculus the action the agent pursues could make sense, however irrational it might seem by its surface features (e.g. climbing over a ladder to get a food item when there is another food item readily available within reach), if we can represent that agents' actions are motivated by a simple utility calculation such that they aim to maximize their expected rewards relatively to the expected costs of their actions. If an agent goes out of her way to obtain a food item, given the availability of other kinds of food items in her close environment, we are going to readily justify her action (however inefficient it seems), given her differential preference for the alternative items. This account, however, holds only if we assume that agents are rational. Without attributing rationality to agents around us, the costs agents expend to get what they want would not be readily represented as an effort that they are incurring subject to their perceptions and beliefs about the state of the environment. Indeed, none of the participants in the studies summarized above inferred that the agent is simply being irrational while interpreting her action. Rather they looked for or inferred the physical constraints in the environment justifying the agent's behavior (e.g. the path being blocked by an inferred obstacle not directly visible for them, as it was the case in Csibra et al., 2003, or by taking into account that the demonstrator's hands were being occupied, as it was the case in Gergely et al., 2002), or they relied on the inferences they drew about the agent's preferences. Previous research has also documented that when an agent acts irrationally and when this behavior is not justifiable by the agent's desires or the physical constraints, adult participants infer that the agent has incorrect beliefs (Baker, Jara-Ettinger, Saxe, & Tenanbaum, 2017) and young children can infer that the agent is ignorant (e.g. that the agent does not know about the presence of a closer by food item, and so she takes the more effortful path to get another food item of the same kind, Jara-Ettinger, Floyd, Tenanbaum, & Schulz, 2017).

Action understanding is structured around an explicit assumption that agents are rational, and when agents do not appear to act efficiently, we can readily form a wide range of inferences that go beyond the here-and-now in order to understand underlying causes of their behavior. Being able to consider that agents act for reasons and might have several goals in mind while acting allows us to interpret seemingly arbitrary actions of others. For example, when an agent's actions were inefficient (and if the given goal failed to explain her behavior), adult participants judged the agent as completing a sub-goal instead (Baker, Saxe, & Tenanbaum, 2009). Another study that used no external physical goal in the display has also found that adult participants judged the agent performing inefficient movements (in comparison to straight movements following the same trajectory) as engaging in communicative displays, controlling for the visual complexity of the inefficient movements (Royka, Aboody, & Jara-Ettinger, 2018; also see Schachner & Carey, 2013).

Critically, in communicative behaviors (e.g. nodding, winking, Royka et al., 2018) or in non-communicative body movements (e.g. dancing, Schachner & Carey, 2013) there is no visible goal one could attach to these actions (as in the case of eating an apple). Even though goal representation is seen as equivalent to representing external end states, not every intentional action event serves a perceptually identifiable goal. Hence when there is no visible external goal that could be attached to the action, the performance of the action can be conceptualized as being the goal. Similarly, when there is a potential external goal the action could be related to, but when the actions do not relate to it efficiently but still were performed intentionally, the actions themselves could become interpreted as the sub-goals that provide normative information as to how the action ought to be performed (as in the case of conventional and ritualistic behaviors).

1.1.2 Interpreting actions beyond physical efficiency: the role of social factors

Not every behavior has a readily observable external goal, and goals have a hierarchical structure (e.g. reaching for the bread is an [enabling] sub-goal to achieve the higher-level goal to make a sandwich, which in turn contributes to an even higher-level goal of satisfying hunger). If we were to make sense of actions only when external goals are physically apparent or without being able to interpret the nested structure of sub-goal-goal in hierarchies, then we would be constrained by a primitive action understanding that only applies to actions directed to create a physical change in the environment. However, human culture has a wide range of rituals, conventions and norms, and we practice those faithfully and recognize them easily as serving a higher goal that is often times not physically apparent in the environment (e.g. taking the hat off to show respect in the church). In a similar vein, given our propensity to explain actions in terms of their goals (Csibra & Gergely, 2007), actions that appear inefficient still warrant an interpretation. This is especially so when the immediate physical context and properties of the agent fall short in justifying seemingly irrational actions. In line with previous work (Gergely & Csibra, 2006; Gergely & Jacob, 2002; Kiraly, Csibra, & Gergely, 2013), I propose this interpretation could be social in nature. Hence in the next section, I will focus on socially stipulated actions, such as ritualistic actions, with the argument that rationality in action is not always constrained by instrumental physical efficiency, or agents' subjective desires or their knowledge states (i.e. ignorance), but also by the social context in which they are performed.

Imagine an agent approaching a food source that she has a liking for. Imagine also that she colored her face wearing the special amulets of her tribe as she moves along the path that will take her to her goal by performing dance-like movements. To our adult intuition, it is obvious that this is how she goes for a hunt in certain occasions. We do not interpret the action anymore within an instrumental stance, but take into consideration a certain variety of socially relevant behavioral markers and displays the agent presents during her action (e.g., the pattern of paint she puts on her face, her stylistic behavioral manner of action delivery, or the traditional ornaments she wears), and interpret the action in terms of the framework of "ritual stance" (Hermann, Legare, Harris, & Whitehouse, 2013; Legare & Hermann, 2013; Legare & Souza, 2012, 2014; Nielsen, Kapitany, & Elkins, 2015). Ritualistic actions are prescriptive in nature (Staal, 1975); unlike every-day instrumental actions covered extensively above, they are not "interchangeable and auxiliary" (Schachner & Carey, 2013, p. 310). After all, there can be several ways to attain a goal efficiently (while preparing food for yourself you can put the spread inside the bread, or on top; both will equally fulfill the goal if you would like to have a bread with some cheese spread), however, for rituals the aim is to carry out the actions in a precise manner (Bloch, 1974; Staal, 1975), even if they are efficiency-wise arbitrary.

Cultural anthropologists have aimed to decode the meaning of rituals over the years. When they asked the individuals performing them what a ritual means or why it is carried out in a particular way, people were often unable to articulate coherent answers. Most deferred to a tradition saying that it was how their ancestors did it and simply stated that "this is just the way it is done". Despite not being able to justify ritualistic actions, they nevertheless stated that there is a profound significance to these apparently arbitrary actions and conveyed that it is crucial for rituals to be performed in the exact way (Bloch, 2004; also see Staal, 1975). It seems that, within rituals, the actions being performed do not originate from the agent performing them, rather they are stipulated in advance and are not a product of individual innovation but acquired through cultural transmission (Humprhey & Laidlaw, 1994; Legare & Souza, 2012). Not having a physical-causal explanation in relation to the end-state, also known as causal opaqueness (Herrmann Legare, Harris, &, Whitehouse, 2013; Legare & Souza, 2012, 2014), is seen as an essential property of ritualistic actions. This type of causal opaqueness of ritualistic actions could be another reason why practitioners found it difficult to justify why they do what they do without deferring to an unknown authority for the justification of their actions.

Critically, ritualistic actions are also goal-demoted (i.e. "actions are divorced from their usual goals", Boyer & Lienard, 2006, p. 597; also see Boyer & Lienard, 2020; Kapitany & Nielsen, 2015, 2017). Unlike ordinary, every-day actions, in ritualistic actions it is not obvious what goal the actions serve; so, it becomes unclear to the observer why the agent acted the particular way she did. Causal opaqueness and goal-demotion (or goal opaqueness, Legare, Wen, Hermann, & Whitehouse, 2015) are not dissociated in the literature (Kapitany & Nielsen, 2017), probably because causal opaqueness of the action sequence can give rise to goal-demotion (after all, as soon as the causal structure of the action sequence becomes transparent, so could its potential goal). Furthermore, ritual often tends to be both causally opaque and goal demoted (Kapitany & Nielsen, 2017; Nielsen, Tomaselli, & Kapitany, 2018). Thus, the ultimate goal of the action sequence or the factors that constrain and influence the action steps are not accessible to direct observation (Gergely & Csibra, 2016).

Lienard and Boyer (2006) proposed a number of further features that could characterize ritualistic behavior: compulsion (performing the sequence of actions in a particular order), rigidity (performing the actions in the exact way) and redundancy (repeating some actions inside the rituals) (also see Rappaport, 1999). In rituals, a practitioner might move around the target item three times in a clockwise motion while jumping repeatedly or reiterating the same utterance over and over again; one might "wash the instruments that are already clean, one enters rooms to exit them straightaway" (Lienard & Boyer, 2006, p. 816). Such cues of compulsive adherence to the sequence and the actions themselves, in addition to repetition and redundancy in the actions performed, can help the observer interpret the action(s) from a ritual stance. Given also that ritualistic actions lack material efficacy, not enabling the practitioner (or the observer) to attach an everyday goal to the actions, fidelity could only be possible by attending to fine-grained features of the actions (Boyer & Lienard, 2006). In rituals, after all, the "the manner of doing is intrinsic to what is being done" (Rappaport, 1999, p. 38).

According to Legare and her colleagues (Legare & Hermann, 2013), the ritual stance is an interpretive mode one engages in while making sense of the actions they observe. In contrast with the instrumental stance, in which an observer searches for a rationale behind the actions based on physical causation, in ritual stance the observer seeks out for a rationale for actions based on social convention, given ritualistic actions cannot be interpretable from the perspective of physical causation (Legare & Herrmann, 2013; Legare & Wen, 2014). These two modes of reading represent a continuum, and observers can sway around the continuum taking an instrumental or ritual stance based on the properties of the action, and the relevant social cues accompanying the actions. For example observing someone killing a chicken could give rise to instrumental reading of the behavior (as the agent might want to fulfill her hunger) or to a ritualistic reading of the same behavior (as the agent might want to make an offering to a divine-being) (Legare & Hermann, 2013). Apart from the redundancy of some actions in the sequence, the rigidity in the way they are performed, and the compulsion in the order which they are executed, certain cues such as conventional language use, the number of agents performing the action, synchrony in action performance, equivalency of beginning and end states of actions can allow the observer to adopt a ritual stance while interpreting the actions they observe and interpreting the goal behind them (Hermann, Legare, Harris, & Whitehouse, 2013; Legare, Wen, Hermann, & Whitehouse, 2015).

For example, when 3-6-year old children were shown videos of one or more actors acting on a wooden mallet and pegboard with a specific sequence of actions, they were more likely to faithfully imitate the sequence they were shown with when the experimenter had framed the task in a convention oriented language ("she always does it this way") instead of outcome oriented language ("she gets pegs up") beforehand. Furthermore, when children saw two adults in the videos, instead of one, synchronously acting on two identical pegboards and mallets, participants were more likely to faithfully imitate the arbitrary sequence of actions they viewed on the videos. When they were asked to explain why they did it that way, children who were given a conventional task frame and watched the video of two adults, instead of one, revealed an understanding about the prescriptive nature of the actions they viewed in their verbal responses (e.g. "I had to do it the way they did show me") (Hermann, Legare, Harris, & Whitehouse, 2013). This finding, along with other studies exploiting a similar paradigm (Legare, Wen, Hermann, & Whitehouse, 2015; Watson-Jones, Legare, Whitehouse, & Clegg, 2014), constitutes evidence that children can adopt a ritual stance based on several cues: viewing that an action does not change the initial state of affairs (e.g. equivalency of start- and end-states, Legare et al., 2015; Watson-Jones et al., 2014), viewing that an action is performed by various individuals synchronously (Hermann et al., 2013), or being introduced to the action by conventional language (Hermann et al., 2013; Legare et al., 2015).

The role of contextual cues in faithful imitation was also documented with younger children (Carpenter, Call, & Tomasello, 2005). Twelve- and 18-month-old infants were shown two events in which the experimenter pushed a toy mouse across a surface in two distinctive action styles (by making it hop or sliding it playfully). When there was no house present in the scene during the experimenter's demonstration where the toy mouse can go into as the end state of the action, both age groups exactly copied the manner action of the experimenter whereas when there was a house present in the scene infants were more likely to ignore the manner in which it was pushed. Another study with 14-month-old infants documented that when the experimenter demonstrated a particular style of action by which she illuminated a touch sensitive light box (i.e. the experimenter leaned on the box to get her forehead in touch with its surface to light it up) infants were more likely to copy this unusual means they were demonstrated if the experimenter ostensively addressed the infant during her demonstration. When the experimenter's demonstration was stripped of ostensive signals, infants no longer imitated the unusual means but rather used their hands instead to light up the box (Kiraly, Csibra, & Gergely, 2013). These two studies show that certain contexts drive goal attribution differently. The same action in one case was easily interpreted as a goal-directed action, hence infants had no trouble taking the instrumental stance and achieving the goal by the most efficient means available to them (by directly placing the toy mouse into the house, or by operating the light-box by touching it with their hands). Yet in another case, when the physical goal was not apparent (Carpenter et al., 2005) or when the experimenter emphasized that there is a relevant teaching episode taking place for the infant (Kiraly et al., 2013) (see natural pedagogy theory, Csibra & Gergely, 2006; 2009; 2011; Gergely & Csibra, 2005; 2006; also see presumption of relevance Sperber & Wilson, 1986, 1987), infants interpreted the action either as a goal or a sub-goal. Both demonstrations (sliding the toy to make it go into a house and contacting the head to activate a touch sensitive light-box apparatus) involved causally opaque actions, yet contextual information allowed infants to have a flexible reading of the demonstrations, allowing them to prioritize the end goal (i.e. instrumental stance) or to prioritize the way the actions were performed (i.e. ritual stance).

Given that there is evidence for even 12-month-olds' adoption of the ritual stance while interpreting actions they observe, one important question I still have not addressed yet remains: What functions do the ritualistic behaviors serve? Evolutionary anthropology aimed to provide an answer to this question by first articulating a mechanism from which ritualistic actions originates. According to the hazard-precaution system of Lienard and Boyer (2006; Boyer & Lienard, 2006), stereotyped behaviors were commonly performed in our ancestral world in order to circumvent the present or potential danger such as contamination, intrusion by strangers, or social exclusion. Activated by the perceived threats to fitness, the hazard-precaution system results in precautionary behaviors that involve compulsive actions (for example, repeated cleaning or symmetrically separating objects and substances to fight off the hazard of accidentally consuming toxins). Ritualistic actions in this sense are non-intended outcomes of this system, but nevertheless are conceived as byproducts of hazard-precaution machinery, even though their original adaptive functions have disappeared or diminished in the process (Whitehouse, 2012).

Crucially, ritualistic behaviors are costly as they deviate from everyday actions, taking physical and mental effort, and time at best, and can impose critical harm for the practitioners at worst (imagine fire-walking rituals or male initiation ceremonies). According to Henrich (2009), owing to the cost agents incur in performing them, ritualistic actions function as sincere signals of commitment (please see anthropological work on the notion of "rituals as commitment devices", Frank, 1988). They act as credibility enhancing displays which provide evidence of an agent's commitment to the beliefs, preferences, and goals of her group. Ritualistic actions in this framework act as a sincere signal of commitment "not despite their costs, but exactly because of them" (Lienard & Boyer, 2006, p. 818; Sosis, 2005); and they constitute "hard-to-fake" signals (Bulbulia, 2004; Irons, 2001; Sosis, 2000, 2004; Sosis & Bressler, 2003). Sosis and Alcorta (2003, p.265, also see Sosis, 2005) further argue that ritualistic actions facilitate "the expression and reaffirmation of shared beliefs, norms and values" and thus they might be interpreted as a form of communication indicating one's belonging and commitment to a community. Given their socially stipulated properties and causal opaqueness ritualistic actions are also argued to be suited for high fidelity cultural transmission as "they are resistant to individual innovation and change" (Legare & Wen, 2014, p. 9).

Following from these hypotheses, Legare and her colleague argued that the main function of ritualistic actions is to solve adaptive problems associated with group living (Legare & Watson-Jones, 2015; Watson-Jones & Legare, 2016). Recent work has also proposed that ritualistic actions bolster affiliation with group members increasing cohesion through strengthening commitment of group members to collective goals of the group (Whitehouse & Lanman, 2014). The limited amount of available empirical research investigating the functions of ritualistic actions shows that engaging in ritualistic acts increases group affiliation among preschool aged children. For example, Wen and her colleagues (Wen, Herrman, & Legare, 2016) found that when 4-6-year-old children participated in a scripted synchronous necklace making task with their own minimal group members over the course of two weeks, they were more likely to show preference for their own group than the children who participated in a similar activity that did not have ritual-like characteristics.

Several authors also discussed that ritualistic actions act as reliable markers of group membership (as participating in group-specific conventional activities allows one to identify her ingroup members, Watson-Jones & Legare, 2016; Legare & Harris, 2016). However, there is no direct empirical evidence supporting this claim. While I agree with the role of rituals in identifying group members and boosting affiliative relations with groups members, in thesis I argue that representing social groups facilitates adopting a *ritual stance*. Characterizing others as belonging to different social groups could help infants and young children socially situate the actions they observe, however irrational they might seem by their surface properties, as group specific conventions, which should be acquired and transmitted as they are, with high-fidelity (Gergely & Csibra, 2006), and should not be generalized to any other social group's cultural repertoire. This ability to distinguish social categories helps young children, cultural novices, to determine what kind of behaviors they should perform to operate within a group and to coordinate their behavior in maneuvering their way around in group living, as they need to learn and adhere to the cultural practices of their own group.

But what is exactly a social group and which motivations give rise to group cognition? The sub-section below aims to elucidate the nature of social group cognition by providing a framework about why humans are prone to organize their social world in terms of social categories, with guiding the reader through several accounts in explaining children's social category understanding. I will argue that social group cues give information about the shared knowledge repertoire of informants, which includes a vast body of socially stipulated conventional practices such as ritualistic actions, as covered in this sub-section. Only by representing social groups can cultural novices track culturally competent sources to learn causally opaque actions from and attribute the knowledge of the same kind to other parties belonging to the same social group. Henceforth, I use the term "cultural knowledge" in this thesis to refer to causally opaque, costly (as they deviate from efficiency) conventional action forms shared among the members of a social group, whose utility could be rather traced to the benefits associated with group living, thus justifying the costs they require from the agents performing them.

1.2 What are social groups (and what are they good for)?

Seminal work in social psychology in the 1970's documented that humans display proneness to divide the social world into "us" versus "them", even when given minimal and arbitrary cues that bear no direct relevance to real life social categories such as gender, race, religion, ethnicity, or nationality (Billig & Tajfel, 1973; Tajfel, 1970; Tajfel & Turner, 1979). When the researchers randomly divided the adult participants (e.g. based on a coin toss or drawing preference) into two groups using an arbitrary t-shirt color cue, they found that participants displayed what is called "ingroup favoritism": they allocated more resources to their own minimal group than the other group. Young children also display sensitivity to minimal group membership (Dunham, Baron, & Carey, 2011; Plotner, Over, Carpenter, & Tomasello, 2015; Oostenbroek & Over, 2015; Richter, Over, & Dunham, 2016; Sparks, Schinkel, & Moore, 2017). When 5-year-olds were assigned to novel color groups (Dunham, Baron, & Carey, 2011), they were more likely to favor their own minimal group upon being asked to indicate their liking for the target individuals belonging to the same colored t-shirt group as them in comparison to those with a different colored t-shirt. They were also more likely to share with their own minimal group members in a resource allocation task, and to form positive implicit associations with their ingroup members. Recently it was documented that even 3-year-olds show more liking for their own minimal group that they were just randomly allocated in based on arbitrary and meaningless cues if given explicit group labels (Richter, Over, & Dunham, 2016).

Minimal group paradigms have proved to be useful tools in eliciting group belongingness even in young age groups and gave researchers a platform to study intergroup cognition without such confounds as familiarity and stereotypes about already existing social groups (Richter et al., 2016). However, minimal groups (especially in the absence of group labels) might not be likely to elicit enduring representations about different groups as there is no real reason for participants to extend their ingroup favoritism to other individuals who wear the same colored t-shirt as themselves long after the experiment is over. After all, t-shirt color is an arbitrary cue that one can easily get rid of unlike one's gender, race or native language.

Such social categories as gender, race and language are associated with impenetrable boundaries that make it difficult for their members to leave the group (or for others to enter) and they have larger group sizes and long histories (Campbell, 1958; also see Lickel, Hamilton, Wieczorkowska, Lewis, Sherman, & Uhles, 2000). From a very early age on humans show sensitivity to these social categories, as reflected by their differential visual attention to social category distinctions (Quinn, Yahr, Kuhn, Slater, & Pascalis, 2002; Bar-Haim, Ziv, Lamy, & Hodes, 2006; Kinzler, Dupoux, & Spelke, 2007). Furthermore, even young children are found to essentialize certain social categories, construing social category membership as something stable with its members sharing similar psychological traits (psychological essentialism: Hirschfeld, 1996; Heyman & Gelman, 2000; Gelman, 2003; Rhodes & Gelman, 2009; Diesendruck & haLevi, 2006; Gil-White, 2001; Diesendruck, Goldfein-Elbaz, Rhodes, Gelman, & Neumark, 2013; Rhodes, Leslie, & Tworek, 2012). Studies documenting essentialism in young children, especially for social categories that are culturally relevant to them (e.g. ethnic essentialism among young Israeli children, Birnbaum, Deeb, Segall, Ben-Eliyahu, & Diesendruck, 2010), suggest how prone we are to categorize the social world around us and draw generalizations however limited our *exemplars* are. Yet being able to do so, despite many negative social ramifications, could in fact be an adaptive trait that helps us 1) fulfill our need to belong by establishing positive social interactions with similar others (i.e. affiliative goals; Baumeister & Leary, 1995; Nielsen & Blank, 2011), 2) track social relations and identify agents with whom we should form alliances (i.e. coalitionary goals; Kurzban et al., 2011; also see Jin & Baillargeon, 2017) and 3) identify epistemic sources who are good candidates from whom we can learn the culturally relevant knowledge of the social group we live in together (i.e. learning goals; Olah, Elekes, Brody, & Kiraly, 2014; Soley, 2019).

Before inviting the reader to evaluate the evidence corroborating the perspective driven by the learning function of social categorization for this thesis, below I review several approaches offering different functions for infants and young children's early sensitivity for the informants belonging to the same group with them. Note, that these accounts are not necessarily mutually exclusive and depending on one's goals in a certain social context (Over & Carpenter, 2012), different views can provide a better explanation in helping us understand why humans, and possibly other species too (yet not as profoundly as humans do [Haun & Over, 2015]), show such proneness to social categorization. Just like any other cognitive process, categorizing and reasoning about social phenomena is an interplay between cognitive and motivational factors and is highly context-dependent (Brown & Pehrson, 2019; Macrae & Bodenhausen, 2000; Turner, Oakes, Haslam & McGarty, 1994). When the social meaning of a category is not relevant, then there might be no need for the activation of that particular social category, let alone for inferences and predictions that follow with that category's activation.

1.2.1 Psychological functions of group cognition

> "*…that part of an individual's self-concept which derives from his membership of a social group (or groups) together with the value and emotional significance attached to that membership*." (Tajfel, 1978, p. 63)

As with categorizing other entities around us (from kitchen appliances to garden tools, from cats to dogs), social categorization also simplifies and brings order to complex social stimuli and structures our world (Tajfel, 1981). According to pioneers of social psychology, social categories, additionally, "provide a system of orientation for self-reference" (Tajfel & Turner, 1979, p. 40) by constructing and characterizing an individual's place in the social world (see social identity theory; Tajfel, 1974). Groups, according to Tajfel and his colleagues (1979), are defined as a collective set of individuals who recognize themselves to be the members of the same social category, and who also have an emotional response in this particular social definition of themselves as belonging to that group. Social identity discourse emphasizes the role of social groups in forming a sense of self by the mere virtue of knowing that one belongs to a certain social group, and not to another. Ingroup favoritism and outgroup hatred (or outgroup derogation; Brewer, 1999) are discussed as consequences of this process according to social identity theory, as individuals who would like to enhance their self-image would favor the members of their group while they would evaluate other groups' members negatively. Gramzow and Gaertner (2015) put forward this hypothesis, more specifically as "the self as an informational base", to explain how self-evaluations link to ingroup favoritism. According to this account, since self-evaluations are often positive for most individuals, the way ingroup memberships are evaluated should be also positive, and there should be a correlation between the extent one evaluates her or himself positively and the tendency to evaluate her or his group favorably. In their account, ingroup favoritism emerges as a consequence of having a positive self-construal. Empirical evidence for this hypothesis comes from studies with adults (Gramzow & Gaertner, 2015; Falk, Heine, & Takemura, 2014; but also see Rutland, 1999).

Despite the evidence for the notion that social groups are central to one's identity and that one behaves in ways to enhance her self-image, leading to ingroup favoritism (Gramzow & Gaertner, 2015), infants show sensitivity to social groups before developing a sense of self and also long before they could be concerned with enhancing their self-image by viewing their ingroup members positively (Jin & Baillargeon, 2017). Developmental work trying to underpin the origins of our motivation to see our group members through a positive light and to judge other group members negatively documented ingroup favoritism with relatively older children

(Bennett, Barrett, Karakozov, Kipiani, Lyons, Pavlenko, & Riazanova, 2004; Buttelmann & Böhm, 2014; Fehr, Bernhard, & Rockenbach; 2008; McLoughlin, Tipper, & Over, 2018; Plötner, Over, Carpenter, & Tomasello, 2015; Renno & Shutts, 2015; Sparks, Schinkel, & Moore, 2017). For example, Buttelmann and Böhm (2014) tested 6- and 8- year old children in a computerized allocation task upon randomly assigning them to minimal groups (i.e. yellow and green groups, determined by a lottery ticket). On one side of the computer screen, children saw a puppet from the same minimal group they had been assigned to, and on the other side another puppet belonging to a different group. Between the puppets stood an empty box the children could click on if they did not want to allocate the target item to any of the puppets. In each allocation trial, the target item appeared on the top of the display and children were instructed to indicate which puppet (with the choice of putting it away to the empty box if they did not want to allocate that item to any puppet) they would like to give the target item to. Only at the age of 8 years, children showed both ingroup favoritism and outgroup derogation: they were more likely to give the target to their ingroup member if the target item was positively valanced, and to the outgroup if the target was negatively valanced (e.g. spider).

While one can still consider that one needs to have a sense of self to recognize that she or he belongs to a social group, and that being a member of a social group helps one to define and revise one's self-concept with the value and emotional significance attached to that membership (Tajfel, 1978), one should also admit that making use of belonging to a group in order to create a sense of self is not a likely function of infant social categorization. Numerous studies depicting ingroup favoritism in children by using allocation paradigms similar to those described above did not provide any compelling evidence that children's positive evaluations of their ingroups stem from their need to boost their self-image. Additionally, non-human animals also have a rudimentary representation of social boundaries (Campbell & de Waal, 2011; Haun & Over, 2015). In light of these, recent developmental work on social group cognition with younger children has rather offered evolutionary perspectives on the functions of differentiating and categorizing our social world.

1.2.2 Evolutionary perspectives

Living in groups and recognizing intergroup boundaries are functional for stable and successful cooperations (Brewer, 1999). Given that group living is characterized by mutual expectation of loyalty, the expectation of ingroup support helps one to build positive attitudes and interactions with other group members encouraging faithfulness to the norms and behaviors of the group she or he lives in, so that she could be identified as a good ingroup member and continue reaping the benefits of the group living (Brewer, 1999). Indeed, ingroup support is one of the basic moral principles guiding our early and universal reasoning about other individuals (Ting, Buyukozer-Dawkins, Stavans, & Baillargeon, 2019; Jin & Baillargeon, 2017; Bian, Sloane, & Baillargeon, 2018; also see Rhodes & Chalik, 2013; Rhodes, 2013; Mich, Over, & Carpenter, 2016 on a similar account that shows how social categories characterize individuals who have intrinsic obligations to one another in young children).

Recent studies documented that 13 to 17 month-old infants expect individuals belonging to the same social group to provide help to each other, but they view helping as optional when the individuals belong to different social groups or when social group cues are removed (Jin & Baillargeon, 2017; also see Ting, He, & Baillargeon, 2019). They are surprised when a harm event (i.e. one agent picking up a stick and jumping on the other one) occurs between two animated agents belonging to the same group (i.e. two blue squares) in contrast to between two agents who belong to different groups (i.e. a circle and a square) (Jin & Baillargeon, manuscript in preparation). In the context of resource distribution, when resources are scarce, infants expect the ingroup distributor to share only with the ingroup members rather than expecting the distributor to divide the items equally between both ingroup and outgroup individuals (Bian, Sloane, & Baillargeon, 2018; also see equity-based principle of fairness).

Numerous findings documenting the evidence for infants' expectations for ingroup support go beyond showing that infants and young children favor their group members in their own interactions as they additionally demonstrate that infants can extend their moral reasoning to guide their expectations about how other in- (and out-) group members should behave. The basic principle of ingroup support (and also ingroup loyalty, see Ting, Buyukozer-Dawkins, Stavans, & Baillargeon, 2019 as a concept defining our early tendency to interact with individuals belonging to our ingroup and to adhere to their object choices and behaviors) guides infants' early reasoning about social group members even when they are not part of those particular social groups. This is the case even when the individuals the infants view on the screen bear no resemblance to themselves (Powell & Spelke, 2013; Bian et al., 2018) as they are animated shapes with agency attributes or hand-played puppets. Apparently, the principle of similarity is a reliable cue for infants, helping them recognize and cluster social agents into different groups. There are indeed reasons why the cue of similarity can be the dimension humans have used throughout evolutionary history to recognize and differentiate social groups: as these two being not mutually exclusive, 1) similarity is a proxy for familiarity; 2) similarity is a reliable cue marking coalitional interactions.

1.2.2.1 From kin recognition to homophily-based accounts to coalition tracking system

Kin recognition governs social behavior in human and non-human animals in the domains of altruism, conflict and mating (see kin selection theory, Hamilton, 1964). Hence one can think of kin as the first and foremost ingroup that every animal belongs to and distinguishes from non-kin through several mechanisms (Liberman, Oum, & Kurzban, 2008). One possible mechanism that individuals use to recognize kin is governed by the principle of familiarity and prior association (Blaustein, 1983). In small groups, every individual can form some level of familiarity with one another and hence relatives can recognize each other if they predictably met each other before. However, in our ancestral world, humans did not live in small social groups exclusively consisting of individuals who are genetically related to one another (and for good reasons, see Liberman et al., 2008 on evolved psychological mechanisms designed to avoid kin as sexual partners to prevent biological consequences of inbreeding). Also, as group size increased at one point in evolutionary history, the frequency with which group members interacted with each other decreased. Thus familiarity, albeit still important, cannot be the only factor aiding kinship recognition, let alone guiding individuals to assess with whom they would share their resources, whom they would help, or from whom they would receive information. Instead individuals should guide their interactions with others based on another factor that should reliably correlate with familiarity. One "proxy measure" for familiarity is similarity[2] (Haun & Over, 2015).

Given the high likelihood of members of a social group having similar repertoires of behaviors, norms, and rules of conduct so the social group could be a functional unit, similarity can be one index we are attuned to attend in a myriad of social interactions with others. Homophily-based accounts (Haun & Over, 2015) argue exactly this: humans have preference for others who are similar to themselves in their early interactions and this human-specific reliance on similar others provides a reliable way for socially differentiating others as belonging to our group (and others as not belonging to our group, if not as outgroup) (Mahajan & Wynn, 2012, but also see Cruz-Khalili, Bettencourt, Kohn, Normand, & Schlinger, 2019).

[2] Sociobiological literature also attests to the role of similarity in kin recognition. One mechanism through which kin recognition operates is phenotypic matching. This mechanism assumes phenotypic similarity is correlated with genotypic similarity. Individuals assess similarities and differences between their own phenotypic markers and those of unfamiliar others; which in turn helps them infer their genetic similarity to others, aiding kin recognition (Blaustein, 1983).

Empirical studies document profound evidence (albeit sometimes indirect) for human homophilic preferences: from infants' orientation response to those who look, talk or dress similarly to themselves, to toddler's responses to those who have the same preferences as themselves, to studies showing how children adhere to majority biases forgoing their own behavioral tendencies to be like others (for a review, Haun & Over, 2015). The burgeoning literature on norm psychology and group behaviors reflects on our conceptualization of the social world based on similar others: we do not only selectively attend to similar others, but we strive to be like others (Carpenter & Call, 2009; Uzgiris, 1981; also see "group-mindedness", Tomasello, 2019), possibly to be acknowledged as good ingroup member ensuring that we get the credit we deserve (i.e. reciprocation).

However, according to coalition accounts, natural selection shaped cognitive mechanisms to track coalitions instead, given the need of our ancestors to navigate their social world successfully (Kurzban, Tooby, & Cosmides, 2001). This computational machinery is designed to be responsive to cooperation and competition patterns, and to the cues that might "purposefully or incidentally" predict each individual's alliance (Kurzban et al., 2001). Being sensitive to such cues will allow one to predict unfamiliar others' behaviors helping her or him to negotiate in the social world. In this vein, surely, similarity can be a cue for coalitions, as similarity between individuals and their potential to be allies are covarying observable structures in our social world. Yet we cannot degrade the coalition tracking system to a system that detects and acts on recognizing similarities in the input alone based on co-occurrence of coalitions among similar others. It is a possible, but not efficient way for natural selection to deal with flexible ways that coalitions can be formed when several individuals interact with each other with each having multitude of different characteristics. Furthermore, coalitions might not only be formed by the same group members, but individuals from different groups can form temporary alliances to negotiate conflicts or when they face a common threat (Kurzban, Tooby & Cosmides, 2001; Cosmides et al., 2013).

Given the dynamic structure of coalitions and their transitory characteristics (Cosmides et al., 2013), certain dimensions, however arbitrary they are (take the example of t-shirt color in minimal group paradigms) could become significant for this cognitive machinery as long as they predict coalitional membership. For example, when jersey color was manipulated to be a cue that incidentally predicted coalitions, adult participants' categorization by this arbitrary minimal cue was boosted (Kurzban et al., 2011). Yet not every cue operates by depending on coalition tracking systems. Sex and age, for example, are dedicated dimensions of our social categorization system (Pietraszewski & Schwartz, 2014b), that is designed to reason about social kinds as there are good reasons for our cognitive architectural system to have evolved to track these dimensions since these are highly significant biological constructs that socially guide and motivate our behaviors in several domains. Several studies indeed show how robust our categorization is based on these two dimensions from early on in ontogeny (see Miller, 1983 and Leinbach & Fagot, 1993 on preverbal infants' sensitivity to differences in sex; also see Bahrick, Netto, & Hernandez-Reif, 1998 on 4-month-old's sensitivity to differences in age).

Despite our automatic and spontaneous categorization based on these primary categories of sex and age, there is no social group that is homogenously defined by these two dimensions. Every social group in the animal kingdom requires members of these two social kinds to operate successfully (i.e. a social group require both groups on the dimensions of sex and age to reproduce, Hirschfeld, 1996). Unlike sex and age, language-based social categories mark the boundaries of different social communities since in our ancestral environment social groups were mostly consisted of individuals with homogeneous linguistic repertoires (Baker, 2001; Moya & Henrich, 2016). Language thus serves as a dimension one can reliably rely on to infer someone's belonging to a certain social community (Giles & Billings, 2004; Henrich & Henrich, 2007; Pietraszewski & Schwartz, 2014a; Pietraszewski & Schwartz, 2014b). Critically, the way an individual speaks provides a rich array of information about what sort of person she or he is (Hirschfeld & Gelman, 1997). For all these reasons, the dimension of language as a social category deserves special treatment, as it allows us to make rich inferences about the individuals in several social domains (preferences, skills, knowledge). The section below aims to grant this special treatment to language-based social categories.

1.2.3 The privileged case of linguistic social groups

Even though infants are attuned to differences between novel individuals based on several cues (e.g. race; Kelly, Quinn, Slater, Lee, Ge, & Pascalis, 2007), this does not necessarily mean that recognizing the differences between individuals based on such dimensions would guide their social interactions with others. In fact, young children do not show a strong social preference for *just any* social category. When racial group membership of the agents 5-year-olds viewed was pitted against their accent, children privileged native accent over same-race (Kinzler & Dautel, 2012). This study alone suggests that young children privilege linguistic cues over race information while deciding with whom to socially engage. Infant studies further documented that even 10-month-olds chose to selectively interact with informants speaking the same language with them (Kinzler, Dupoux, & Spelke, 2007) whereas a similar paradigm failed to reveal selectivity in 10-month-old's preferences for the same racial group (Kinzler & Spelke, 2011). When 10-month-old infants were presented with short movies of two people speaking either in the same language or in a foreign language, they preferentially reached for the object offered by the person who spoke in the infants' native language (Kinzler et al., 2007). Yet when 10-month-old Caucasian American infants viewed two individuals on the screen, a white and a black female, each holding an identical toy, infants were at chance level in approaching the same-race individual to obtain the toy from (Kinzler & Spelke 2011). More recent studies further reported the effect of same-language in guiding preferences of young infants on food items (Shutts, Kinzler, McKee, & Spelke, 2009), on artifact functions (Kinzler, Corriveau, & Harris, 2011) and on novel tunes (Soley & Sebastian-Galles, 2015).

Why are infants motivated to socially engage with an agent speaking the same language with them in non-verbal object-choice tasks while they do not manifest a similar social preference for an individual sharing the same race with them? I argue that representing social categories based on linguistic cues can reliably equip us with probabilistic power in inferring others' histories and predicting the likelihood of having common cultural ground with them (see also Pietraszewski & Schwartz, 2014a, 2014b). Race, on the other hand, might not equip us with inferential and predictive power as linguistic cues do, at least not in the same degree, with regards to someone's social history. Additionally, given that race is relatively new psychological construct in evolutionary history (Kurzban, Tooby, & Cosmides, 2001; Lieberman, Oum, & Kurzban, 2008), how could we have a specifically designed cognitive adaptation enabling us to show early sensitivity to racial cues?

In the ancestral environment, hunter-gatherers were not likely to encounter individuals of other races as they were mainly traveling by foot, which rendered their social contact "geographically local" (Cosmides, Tooby, & Kurzban, 2003, p. 174,), and different racial groups were not living in close proximity to each other. However, linguistic variations existed in neighboring areas and between the bands that belong to the same nexus (Bowern, 2010). Thus hunter-gatherers did not have to travel far to establish relations with other individuals who had different linguistic repertoires (Bowern, 2010; Pietraszewski & Schwartz, 2014a). Given the occasional contact with other social groups with different linguistic repertoires in ancestral environments, it is not implausible that there was a selection for cognitive adaptation to preferentially detect and attend to those who speak the way we do. Unlike it is in the case for race, which is argued to be a probabilistic cue to signal coalitional affiliation allowing us to discriminate groups of people with whom we affiliate from those with whom we compete (Kurzban et al., 2001; Cosmides et al., 2003), our early sensitivity to the same linguistic repertoire cannot be explained only by an evolved mechanism to detect and track coalitions. Rather, it seems our early sensitivity for shared linguistic repertoire is a product of a dedicated cognitive machinery, that might as well operate independently from the coalition tracking system (Pietraszewski & Schwartz, 2014a), given the inferential and predictive capacity about others' knowledge repertoires, skills, common goals, and social behaviors that it equips us with over and beyond helping us to track coalitions.

Recent empirical research indeed documents that linguistic cues have privileged status in social category judgements (Pietraszewski & Schwartz, 2014b). In a memory confusion paradigm, adult participants were presented with several statements produced by native-accented and non-native-accented speakers on the screen. Participants were more likely to incorrectly attribute a statement produced by the original speaker to another speaker, who shared the same accent with him, later in the recall test (Pietraszewski & Schwartz, 2014a). The researchers then crossed coalitional cues with accent cue, such that speakers wore differently colored t-shirts and belonged to two different charity groups while speaking in the native accent of the participants or in a non-native accent. Despite accent not being predictive of the coalitional cue (i.e. two people can speak with the native-accent but belong to two different charity groups nonetheless), participants still made attribution errors based on the accent cue at similar rates that were found in the original study (Pietraszewski & Schwartz, 2014b). In a follow-up study, when coalitional cues were crossed with race, instead of accent, so that target speakers were white and black individuals, attribution errors based on the race information diminished greatly in comparison to a condition where there was no crossing of race with coalitional information. Coalitional stimuli, when not predicted by race, reduced categorization by race as these studies documented. However, the same coalitional stimuli did not have any effect on the categorization adult participants made based on the accent, showing how robust language based social categorization is.

When it comes to how social categorization driven by linguistic repertoire differences of others influences our social judgements, previous research documented not only that adult participants remember the statements made by foreign-accented speakers less accurately but they also evaluated them as being less credible than the statements made by native-accented speakers (Lev-Ari & Keysar, 2010, 2012). However, when there was a mismatch between the way the target person spoke and the way she or he appeared, linguistic cues seem to prevail yielding positive social judgments. For example, in a recent study when Turkish candidates spoke with a native-like German accent in a job interview, they were evaluated positively because of their native-accent and were more likely to be hired than German looking native accented candidates (Hansen, Rakic, & Steffans, 2018) showing the privileged role of linguistic cues in enabling us to form enriched inferences about others (i.e. it takes a lot of effort to reach native-like accent proficiency after all, which might reflect on the perception of the candidates as being hard working).

In addition to work with adults, studies with young children suggest that the human mind is indeed designed to attend to differences in linguistic repertoires of others in their social evaluations. Dejesus and her colleagues (2018) asked two age groups, 5-6-year-olds and 9-10-year-olds, to judge target individuals' national group membership. Stimuli were photographs of white or Asian adults accompanied by voice clips recorded in English, Korean-accented English, Korean, and an unfamiliar foreign language. In the youngest age group, both monolingual white American children and bilingual Korean American children were more likely to judge English speaking people as "American" and Korean speaking people as "Korean" independently of the racial information given. Only the older age group was able to incorporate racial information in their judgements, evaluating Korean-accented English-speaking targets as American but only if the targets were white. Younger children robustly used the linguistic information while deciding on the nationality of the targets and dismissed the race cue. Nationality judgements of young children, although emerging late (Davies, 1968), hint that even at the age of 5-6 years children use shared language as a proxy for assessing whether the target individual is a member of the common culture.

Soley and Aldan (2020) further showed that 5-6-year-olds evaluated the individuals who speak the same language, but not of the same gender, as having the same cultural knowledge repertoire. Children viewed the photographs of three children: the target child's photograph appeared on the top of the screen and the two photographs depicted two other children below, one of whom spoke in the same language with the target child but was of an opposite gender whereas the other spoke an unfamiliar foreign language but was the same gender as the target child. In the shared knowledge condition, children were told that the target child listened a song that the experimenter played, and that she *knew* the song. In the shared preference condition, children were told that the target child *liked* the song that the experimenter played. In both conditions, participants were then asked to indicate which of the two children depicted below also know or also like the song that the target child heard. Participants in the shared language condition were more likely to choose the child who spoke the same language of the target child as the one who also knows the song. Gender did not influence their choices. In the shared preference condition, children's responses were random.

Despite the privileged treatment of language by preschool aged children in inferring the knowledge state of others in comparison with gender, one can imagine scenarios where gender alone can guide perhaps slightly older children's evaluations about others' knowledge repertoires. For example when the target information taps into a gender-related norms (e.g. I might not expect my mother to know how to fix the broken washing machine at home, as I might not expect my father to know how to knit), preschool aged children could be also using the cue of gender to infer knowledge repertoire of others (Bussey & Bandura, 1984 showing children imitate same-sex models as comparison to opposite sex models). However, given that gender-related norm understanding is a quite late acquisition enriched with norm enforcing social practices, it is no surprise that the link between knowledge and gender is not as robust as the link between knowledge and language, especially when it comes to idiosyncratic forms of knowledge as in the case of songs (Soley & Aldan, 2020; Soley, 2019). Cultural knowledge on the other hand, from artifact use to social norms, needs to be mastered quite early in life. Hence detecting and tracking others who have relevant knowledge repertoire on cultural forms comes with its own epistemic advantage. In this thesis I propose that language has such a privileged status as a social category because it is a cue of shared cultural knowledge, enabling us to track informants who could be culturally credible sources to rely on while acquiring cultural knowledge of the social group we live in.

1.2.3.1 Language as a proximal cue for shared cultural knowledge

Most forms of cultural knowledge are cognitively opaque and cannot be obtained through individual trial and error learning (Csibra & Gergely, 2009; Gergely & Csibra, 2006). Indeed, cultural knowledge is transmitted through direct contact with individuals of the social group the naïve learner lives in (Soley, 2019). Hence it is no surprise that children expect cultural information to be shared only between certain individuals who belong to the same cultural group as revealed by the same-language they speak (Soley, 2019) (in contrast to generic facts which children expect to be shared among everyone [Cimpian & Scott, 2012], also see Diesendruck & Markson, 2011). Furthermore, language (unlike other cues of social group membership, such as sex, age or race) signals communicative competence and acts as a reliable tool to transmit and request cultural knowledge.

Even in infancy there is an expectation for two individuals sharing the same cultural knowledge to belong to the same social group as indexed by their affiliative status (Liberman, Kinzler, & Woodward, 2018). Liberman and her colleagues (2018) had 16-month-olds view video segments showing two individuals performing different but similarly unusual means in their goal attainment (i.e. one individual used her forehead to active a touch sensitive light-box apparatus to emit the light effect while the other pressed her elbow on the surface of the box to do the same). After these familiarization scenes, infants watched the affiliation test video portraying the two individuals as they smiled, said ‘hi’ and waved at each other whereas the social disengagement test depicted those two individuals turning away from each other as they crossed their arms frowning. Infants looked longer at the affiliation test video; they were surprised to see these two individuals, who acted on the same object by different means, *affiliate* with each other (Liberman et al., 2018). Despite not being a direct test for the close link between cultural common ground and shared linguistic repertoire, Liberman and her colleagues’ measure of affiliation might be argued to tap into a construct that is over and beyond two individual’s motivation to communicate, as it can additionally reflect those individuals’ capacity to communicate with each other. Nevertheless, this study, albeit indirectly, suggests that infants are sensitive to differences in cultural action repertoires of others and can infer the affiliative relations between individuals based on the way they act.

More direct evidence showing that young children expect common cultural ground to be shared among those who speak the same way comes from a novel eye-tracking paradigm. Olah and her colleagues (2014) manipulated the dimension of conventionality in the way the model used familiar tools, to document that toddlers map foreign-language use with non-conventional tool use. In this study two-year-olds viewed a model depicting different tool-use actions in two between-subject conditions. In the conventional condition, the model used a tool in a regular way (e.g. used a fork to eat food) and in the non-conventional condition, the model used the same tool in a rather unusual way (e.g. used a fork to comb his hair). During the test, toddlers were presented with two photographs: one depicting the model they had already watched acting non-verbally during the familiarization, and the other belonged to an age- and gender- matched unfamiliar adult. The voice of a man speaking in an unfamiliar foreign language was heard from the speakers. Upon hearing the voice, children were found to fixate first on the model than on a novel person in the non-conventional condition upon hearing the foreign language, expecting the model who used the familiar tool in a non-conventional way to speak a foreign language. Children in the conventional condition showed the opposite pattern in their first fixations upon hearing a foreign language: they fixated on the novel person rather than the model. In other words, when the model used the familiar tool in a conventional way, 2-year-olds expected the model not to be the source of the foreign language.

Evidence showing the close link between language and cultural knowledge was also documented in behavioral studies (Buttelmann, Zmyj, Daum, & Carpenter, 2013; Howard, Henderson, Carrazza, & Woodward, 2015). Infants as young as 14-month-olds can guide their learning in acquiring cognitively opaque forms of actions by selectively relying on the informants who speak the same language with them in video displays (Buttelmann et al., 2013). Three-year-olds too rely on their own speakers in a similar paradigm (Howard et al., 2015). Additionally, a recent tool-use paradigm (Olah, Elekes, Peto, Peres, & Kiraly, 2016) found that young children’s faithful copying of the informants speaking the same language with them incites them to commit scale errors, suggesting that children form representations of tools and their functions in rather rigid ways when the tool function was demonstrated by their native-language speaker. Yet when the tool function demonstration was delivered by a foreign language speaker, the rate at which 3-year-olds committed scale errors dropped significantly. This time they flexibly chose the toy that brought about the intended goal more efficiently rather than the one which looked identical to the toy the demonstrator endorsed, but was otherwise too big to achieve the goal with (Olah et al., 2016).

Despite not manipulating the language the demonstrators speak with, studies with minimal groups, further documented that children were more likely to copy the irrelevant actions they were demonstrated when they identified themselves with their minimal group demonstrator (Gruber, Deschenaux, Frick & Clément, 2019) and when minimal cues were manipulated to be socially meaningful for the task (Schleihauf, Pauen, & Hoehl, 2019). Yet the influence of artificially formed minimal groups on the rates of "overimitation" (Lyons, Young, & Keil, 2007) is rather weak as the authors themselves discussed in length (Gruber et al., 2019; Schleihauf et al., 2019). It seems that without additional information that would make minimal groups a meaningful social category, young children do not rely on their group members while learning arbitrary and causally opaque action sequences. However, as children get older, their reliance on minimal groups might start to resemble their reliance on the members of their language speakers. A recent study (Li, Liao, Cheng, & He, 2019) found that 6-year-olds, but not 4-year-olds, were more likely to behaviorally match their responses to their own minimal ingroup members at the expense of efficiency. In Li and her colleagues' paradigm, an ingroup informant pressed a button over another to activate a music box, yet her button choice did not reliably activate the music box with every press (i.e. it only worked 2 out of 4 times). In contrast, an outgroup's choice of the other button produced a reliable effect with every single press. Six-year olds, even if their ingroup informant's button choice did not reliably activate the box, still chose the button their own minimal group member preferred.

Why did three-year-olds in the scale-error paradigm of Olah et al. (2016) and six-year-olds in the music-box paradigm of Li and her colleagues (2019) persisted in choosing the tool or the button selected by their own group members despite the visible inefficacy of both choices bringing about the goal? Why were infants in Buttelmann et al.' study (2013) or older children in imitation paradigms (Gruber et al., 2019; Schleihauf et al., 2019) more likely to imitate the sub-efficient manner actions or action sequences that were irrelevant for goal attainment, when these were demonstrated by informants belonging to their own social group? I propose that this is due to their motivation to acquire culturally relevant knowledge from informants belonging to their own social groups.

Begus and her colleagues (Begus, Gliga, & Southgate, 2016) showed that, indeed, even preverbal infants' preferential reliance on the speakers of their own language is epistemically motivated. In an EEG paradigm, when 11-month-old infants were shown two demonstrators, one speaking in their own language and the other in a foreign language, there was a greater anticipatory theta activity in the infant brain upon viewing a native language speaker indicating their anticipation to receive relevant information. It is highly likely that shared language cue signals the potential of a common cultural background between two parties. That helps the naïve learner to infer they are not only similar in the language they speak, but also in several other socially significant domains that the learner has yet no cultural expertise in. In this vein, shared language can reliably guide naïve learners to the informants who could be the optimal teachers from whom they can learn culturally relevant knowledge. Another premise following from this hypothesis is that same-language cue might act as another ostensive signal for children, instantiating biases that enable them to form generic and enduring representations about the content of the knowledge transmitted by others who speak the same way as they do. Hence sensitivity to same-language cue might reflect, as ostensive signals do, a dedicated adaptation to ensure optimal cultural transmission. Despite not being the central topic of this thesis, I will come to the discussion of same-linguistic cue as an ostensive signal in the general discussion. For now, I restrain the topic to the influence of same-language – and same-social group cues in infants and young children's learning and attribution of cultural forms of knowledge.

1.3 Overall aim and overview of the empirical studies

Given the corroborating evidence on the link between common cultural ground and shared language, in addition to behavioral studies showing infants' and young children's reliance to the informants belonging to the same social group with them, and sometimes atthe expense of efficacy, I propose that our early propensity to interact with and learn from the individuals sharing the same-language repertoire with us is epistemically motivated. Naïve learners' reliance on individuals sharing the same language with them gives them a learning advantage when it comes to acquiring cultural knowledge of the social group they live in, boosting their social fitness in the long run and eliminating the exhaustive, and in the case of cultural knowledge, rather implausible process that comes with trial-and-error learning and exploration in the short run. Given also different cultural repertoires each social group has, naïve learner's propensity to track and conform to the informants sharing the same language, and hence the same cultural ground with them, enables them to be epistemically vigilant. It is after all, not adaptive to learn idiosyncratic cultural forms of any social group if they have no relevance in the cultural group they are socializing into.

Here I operationalize cultural knowledge as conventional, causally opaque action forms following the work of Legare and her colleagues on ritual stance (Legare & Hermann, 2013; Herrmann Legare, Harris, Whitehouse, 2013; also see Kapitany & Nielsen, 2005; Wilks, Kapitany, & Nielsen, 2016). I have organized the empirical studies included in this thesis into five different chapters. Each empirical chapter is motivated by the aim described above and explores related but different questions driven by the role of social group cues in aiding infants and young children to learn about causally opaque actions. Throughout the empirical work carried out for this thesis, infants or young children were presented with sub-efficient manner actions (actions that are sufficient to bring about the end state but are not the most efficient actions available in the physical environment, note that in Chapter 4 I labelled this as inefficient action). Given the cognitive opaqueness of sub-efficient manner actions (it is not causally apparent from the physical environment alone why the agent would intentionally carry out that particular sub-efficient action, but not a different one) I have focused on sub-efficient object directed actions with the argument that they are good candidates for conventional actions, henceforth for cultural knowledge.

Chapter 2 explores whether 18-month-old infants selectively reenact sub-efficient actions demonstrated by an informant speaking the same language as the infants over an informant speaking a foreign language in an imitation paradigm. Buttelmann and his colleagues (Buttelmann, Zmyj, Daum, & Carpenter, 2013) investigated 14-month-old infants' selective imitation of a linguistic ingroup over a linguistic outgroup demonstrator in a paradigm where the demonstrations were presented to the infants via video. Another study investigating the same phenomenon (Howard, Henderson, Carrazza, & Woodward, 2015) failed to find a similar pattern of selective imitation among 19-month-old infants when the demonstrations were delivered live. Chapter 2 hence provides a replication of infants' selective imitation from their linguistic group members that closely follows the work of Buttelmann et al. (2013), however with an age group of the above disputing study (18-month-olds), and with a live head-touch imitation paradigm.

Chapter 3 investigates whether 14- and 18-month-olds can represent a particular sub-efficient action as a normative manner guiding the way a social group member *ought to* act on a novel apparatus in a looking time paradigm. Infants were first presented with an ostensive informant speaking in the native language of the infants demonstrating a sub-efficient action on a novel apparatus. Then infants saw two agents entering the scene either speaking the same language as the infants (and hence as the demonstrator) or speaking in a different language (between-subjects). Later infants watched these two agents acting on the apparatus with two novel means: one agent's mean was similar to the demonstrator's, and the other agent's mean was different. This study explores whether or not infants expect how novel agents should act on the apparatus based on the language they are revealed as speaking.

Both Chapter 2 and Chapter 3 exploit language-based social categories in 2) infants' reliance on their own group member in learning sub-efficient actions and 3) their attribution of shared cultural knowledge repertoire to other agents based on the agents' social groups, respectively. Chapter 4, in contrast, exploits cues indicating similarity (i.e. color, shape) and interaction patterns (i.e. proximity, synchronous dancing) between animated characters in infants' social category representations based on the work of Powell and Spelke (2013). In a violation of expectation paradigm, Chapter 4 looks at whether infants expect animated agents to share the movements of their group members in goal-directed actions. This chapter further elucidates on infants' sensitivity and expectation of efficient actions, and whether the expectation of efficiency interacts with the expectation of shared movement repertoires of agents in social groups. Chapter 5 brings the focus to the processes of faithful imitation in cultural knowledge transmission among preschool aged children. It investigates how children integrate sub-efficient and efficient variants of novel means actions as a function of the demonstrator's language. Chapter 6 provides a summary of the empirical work I have undertaken, in addition to giving a general framework that discusses different theoretical accounts on action understanding, imitation and social categories. I will close the final chapter by laying out the future directions this body of work can give rise to.

Chapter II. Eighteen-month-olds' selective imitation of demonstrators who speak the same language as the infant's social community

2.1 Introduction

Human infants can learn a large amount of skills from others even without directly seeing the end-states of the observed actions by the age of 18-months (Meltzoff, 1995). From early on they can also distinguish accidental from intentional actions as documented by their selectivity in reenacting only the actions that were performed intentionally by the model (Carpenter, Akhtar, & Tomasello, 1998). In the "rational imitation" paradigm (Gergely, Bekkering, & Kiraly, 2002) they show sensitivity to the situational constraints of the model (whose hands are either occupied or free when she lights up a lamp by touching it with their forehead) and selectively modulate the manner in which they reproduce the demonstrated goal as a function of their own physical constraints. This is shown by the fact that when the demonstrator's hands are occupied but the infants' hands are free, they don't imitate the demonstrated head-touch action but light up the lamp by using their (free) hand (i.e., they emulate the goal) (see also Buttelmann, Carpenter, Call, & Tomasello, 2008; Király, 2009a; Király, Csibra, & Gergely, 2013; Schwier, van Maanen, Carpenter, & Tomasello, 2006; Zmyj, Daum, & Aschersleben, 2009). As all these studies suggest infants do not blindly copy what they see others doing, but rather engage in a context-sensitive inferential reading of others' action demonstrations on novel artifacts. Yet how do they know which aspects of the action demonstration to focus on while it is (probably) their first time observing that particular action on the novel artifact?

According to the teleological stance (Gergely et al., 2002; Gergely & Csibra, 2003; Kiraly et al., 2013) when we observe an action, we seek to establish a relationship between three relevant elements of the action-event, namely, the action itself, the physical constraints of the situation in which the action takes place, and the goal-state achieved. We only make sense of a behavioral act as a goal-directed instrumental action if these three elements are rationally linked together. For example, infants observing an agent lighting up a touch-sensitive lightbox (*goal*) can make sense of her unusual head-touch action on the box (*action*) as an instrumental goal-directed action if the agent's hands were occupied while performing the head-action (*physical constraints*). This framework does not only allow us to make predictions about unseen goals given the available information about the action and the physical context but also allows us to rationalize or justify why the goal was achieved by the particular manner of means action performed by the agent observed. For example, if I see someone opening the door by using her elbow, I would assume that her hands must be occupied even if I cannot directly see them, given that I know her goal is to leave the room. Even very young infants can infer that seemingly unusual actions are justifiable by the physical constraints in the environment. Csibra et al. (2013) presented scenes to 12-month-old infants in which an agent repeatedly jumped over an occluded area to reach its target on the other side. In the subsequent test events when the occluder was removed, infants expected to see an obstacle in the middle of the path and were surprised if there was no obstacle revealed that would have justified the jumping action of the agent as an efficient goal-approach.

Yet we do not only observe instrumental actions that achieve clearly observable goals by creating a change in the physical environment. We also see other people around us performing customary, normative and sub-efficient actions. For example, your neighbor takes his hat off while entering the church, your butcher performs the slaughter in a very ritualistic manner if he is a Muslim, your mom's meatball specialty includes a very intricate way of preparing the meat (Gergely & Csibra, 2006). Humans are unique in their capacity to acquire and faithfully transmit such cultural action routines, but if a naïve observer would apply the principle of rationality when trying to interpret them, would fail and would discard such actions as non-rational. When it faces various conventional, traditional, or ritualistic actions, the teleological stance of action understanding cannot interpret such sub-efficient cultural actions as rational goal-directed actions. Oftentimes, a conventional action does not create a physical change in the environment, or even if it does the causal relation between the action and the goal state is not apparent (and it might be as well non-existent). There might be an overarching goal behind the conventional action that is not readily observable, which one can successfully realize only if she or he is to interpret the action prescriptively, i.e. as a demonstration of "this is how it should be done". For example, what it looks like a goal to the observer (e.g. taking the hat off) could in reality be a sub-goal that serves the overarching goal of the convention (e.g. showing respect). Also, critically, when it comes to cultural actions, there is not much room for emulation (or for innovation), but rather it is faithful imitation what appears necessary for its successful transmission over generations (Gergely & Csibra, 2006; Legare & Watson-Jones, 2015; Watson-Jones & Legare, 2016). Given their opaque features, what do infants have at their disposal preparing them to readily acquire such cultural actions?

Here we argue that naïve learners must rely on social cues to guide them in acquiring such cultural knowledge. As shown by previous research, ostensive communicative signals (such as establishing eye contact and being addressed by infant/child directed speech or "motherese") help infants to acquire cognitively opaque information (Csibra & Gergely, 2006, 2009; Gergely, Egyed & Kiraly, 2007). They facilitate infants' acquisition of cultural knowledge about the functions of novel tools despite the cognitive opacity of the manner of operation and the inherent function served by the demonstrated novel artefact (Futo, Teglas, Csibra, & Gergely, 2010). Ostensive signals help infants to form novel means-actions representations (Hernik & Csibra, 2015), highlight the hidden dispositional properties of objects (Kovacs, Teglas, Gergely, & Csibra, 2017), and contribute to their learning of seemingly arbitrary sub-efficient actions (Brugger, Lariviere, Mumme, & Bushnell, 2007; Buchsbaum, Gopnik, Griffiths, & Shafto, 2011; Kiraly et al., 2013; Hoehl, Zettersten, Schleihauf, Grätz, & Pauen, 2014; Kupán, Király, Kupán, Krekó, Miklósi, & Topál, 2017; but see also Nielsen, Moore, & Mohammedally, 2012; Shimpi, Akhtar, & Moore, 2013). Ostensive signals achieve this by conveying to the addressee the demonstrator's communicative intention to transfer relevant and new knowledge. They induce a "presumption of relevance" (Sperber & Wilson, 1986) and basic epistemic trust in their addressee (Gergely, Egyed, & Kiraly, 2007), leading the addressee to interpret the informative content of the agent's communicative demonstration as conveying relevant generic knowledge, extending across different agents, situations, and going beyond *here-and-now* (Csibra & Gergely, 2006, 2009; Csibra, 2010).

However, relying on someone's communicative intention alone might not guarantee that the informative content of the ostensive communication is the most relevant piece of knowledge for a novice to acquire. Infants in fact show readiness to pick up signals indicating epistemic unreliability of the informants as indicated by their incompetence (Zmyj, Buttelmann, Carpenter, & Daum, 2010; Stenberg, 2013), inaccuracy (Poulin-Dubois, Brooker, & Polonia, 2011; Brooker & Poulin-Dubois, 2013) or uncertainty (Birch, Akmal, & Frampton, 2010). Moreover, informants might have epistemic constraints especially in the domain of cultural knowledge due to their lack of expertise in the conventional practices of the cultural group they are not a member of. Hence observing an agent showing communicative intention (by using ostensive signals) might not always warrant that the content of her informative intention is relevant to one's learning goals in any context. Therefore, any epistemically vigilant addressee should use more direct cues of relevance, if available, to judge whether the information presented to them is relevant for their own learning goals from early on in ontogeny.

One direct social cue signaling that the information is likely to be culturally relevant and hence worth acquiring is the language the informant speaks. Previous studies showed that by the age of six months infants prefer and learn from the speakers of their own language (Kinzler, Dupoux, & Spelke, 2007; Begus et al. 2016; Soley & Sebastian-Galles, 2015; Marno, Guellai, Vidal, Franzoi, Nespor, & Mehler, 2016). Despite still being under debate (Liberman, Woodward, & Kinzler, 2017, Kinzler & Liberman, 2017; Begus, Gliga, & Southgate, 2016) one function of this early sensitivity is argued to be to direct infants' attention towards the informant who might potentially provide them with the most relevant information (Begus et al., 2016). In an EEG paradigm, Begus and her colleagues (2016) presented scenes to 11-month-old infants with two agents. During the familiarization, one agent provided labels and demonstrated the functions of the familiar objects to infants, whereas the other agent did not provide any label or function information about the objects but rather commented on the object by saying "ooh" or simply held the object. In the test phase, the two agents were presented individually with novel objects in front of them. Theta oscillations (generally regarded as associated with an expectation of relevant information to be acquired, see Guderian, Schott, Richardson-Klavehn, & Duzel, 2009,) were measured during the period before the agents informatively or non-informatively interacted with the novel objects. EEG measures indicated higher theta activity for the informative agent. In a subsequent study, when infants were this time presented with two agents, one speaking in the infants' own language and the other in a foreign language, there was again greater theta activity in the infant brain upon viewing the native language speaker during the test phase, indicating infants' active preparation to learn relevant information from the agent speaking the same language as their own social community. Infants' expectation of relevant information from the agents sharing the same language with them is only documented in a single behavioral paradigm showing 14-month-old infants' preferential learning from televised demonstrators sharing the same language as theirs (Buttelmann, Zmyj, Daum, & Carpenter, 2013; but also see Howard, Henderson, Carrazza, & Woodward, 2015).

Here we hypothesize that infants readily learn sub-efficient actions ostensively demonstrated to them only by the models speaking their own language, even if this demonstration does not present an efficient action means that they could discover by themselves. We argue that this early propensity to rely on the same-language speakers helps naïve learners to acquire cognitively opaque forms of social practices, conventions, and customs of their cultural group. According to this proposal, infants are not *rational* in their imitative behaviors just because they can take the situational constraints of the informants into account (rational imitation: Gergely et al., 2002 ; Kiraly et al., 2013) but also because they can rely on different social cues while interpreting the epistemic constraints of the demonstrators in guiding their learning experiences. Speaking an unfamiliar foreign language is one example of such an epistemic cue that can signal the infant that the speaker may be culturally uninformed about the common social practices of the infants' social group. If infants indeed interpret the foreign-language speaking demonstrator as an unreliable epistemic informant, given that non-shared foreign language the informant speaks signals that she or he is not a culturally competent epistemic source of relevant information, then infants will not be readily show faithful imitation upon viewing the ostensive demonstration of a sub-efficient action by this agent. In contrast, we predict that infants will be more likely to show high fidelity imitation of the sub-efficient action if it is delivered by a demonstrator speaking the same language as their social community, despite the availability of a self-discovered efficient alternative. If the language the informant speaks brings no additional interpretation of being a reliable source of culturally relevant information, infants should not be differentially learning from ostensive demonstrators speaking the same versus a foreign language.

To test this hypothesis, we used Gergely et al.'s (2002) head-touch imitation paradigm. Infants first either viewed a demonstrator telling a story in their own language or in a foreign language. Then the same demonstrator (whose hands were free) performed a head-touch action three times to operate a touch sensitive lightbox by using non-verbal ostensive signals, and she left the room leaving infants alone with the target apparatus. A third group of infants participated in another condition where there was no demonstration of the head-touch action. Our procedure was very similar to the study by Buttelmann and his colleagues (2013) with the exception that our protocol was carried out live while Buttelmann et al.'s was based on video-demonstrations and we targeted 18-month-old infants (an age where infants also show rational imitation; Gellen & Buttelmann, 2019).

We investigated the reenactment of the head-touch action by the infants and registered whether their reenactment of the modeled head-touch action was preceded by their performance of a hand-action as well. For the condition where the demonstrator spoke the same language as the infant our hypothesis was that more infants will reenact the demonstrated head action, and that they will do so even if they had first successfully brought about the effect by touching the lightbox with their hands. We predicted that infants in the foreign language condition will be less likely to reenact the ostensively demonstrated head-touch action, and that their performance will not differ from the participants' behaviors in the no-demonstration condition. We also predicted that infants in the foreign-language and no demonstration conditions will be more likely to rely on the alternative efficient hand action to bring about the effect.

2.2 Participants

Participants were 48 infants (28 females). Mean age of the infants was 577.23 days (18 months 29 days) (range 549-607 days; 18;0 to19;28, *SD* = 16.8 days). An additional 19 infants were tested but excluded due to being fussy (n = 7 [n = 6 in the foreign-language speaker condition, n = 1 in the same-language speaker condition]), being passive (n = 7 [n = 6 in the foreign-language speaker condition, n = 1 in the same-language speaker condition]), experimenter error n = 1 [in the no-demonstrator condition]), parental interference (n = 3 [in the no-demonstrator condition]) and having a regular weekly exposure to a foreign language (n = 1) [in the foreign-language speaker condition]). All infants in the experimental conditions were monolingual, learning only Hungarian as their native language.

Participants were recruited from a database of parents who volunteered to participate in developmental studies. Each parent gave his or her written informed consent for the study, and the procedure was approved by the United Ethical Review Committee for Research in Psychology, Hungary, and conducted in accordance with the Declaration of Helsinki. Infants were given a toy or other small gift in the end of the study independent of their performance.

2.3 Design

Infants were randomly assigned to three conditions: same-language speaker demonstration (N = 16 [10 females], with the mean age of 577.87 days), foreign-language speaker demonstration (N = 16 [8 females], with the mean age of 573.69 days), and no-demonstration (N = 16 [10 females], with the mean age of 580.12 days).

2.4 Materials

The light box apparatus was a circular lamp with a diameter of 13 cm mounted on a black box with the dimensions of 24 x 16 x 4.5 cm. The lamp lit up when pressed on the surface and remained illuminated until released.

2.5 Procedure

For experimental groups, the procedure consisted of a demonstration and an imitation phase. The demonstrator in the same-language speaker condition was the experimenter who was explaining the study to the parent. Hence to make sure that the participants, independent of the condition they were allocated to, had equal exposure to both demonstrators before the experiment begin, both demonstrators interacted with the infant in a free-play context in their respective languages for approximately five minutes in the common waiting area before the experiment began. For the no-demonstration condition, this free play activity was not carried out.

Upon entering the testing room, infants were seated on a parent's lap in front of a small table. For both same-language and foreign-language speaker conditions, the demonstrator sat down across from the infant, greeted him or her by saying "Hi baby hi, how are you? Now I am going to tell you a story, it goes like this". Depending on which condition the infant was assigned to, she or he either viewed the demonstrator's greeting and story in her own language (i.e. Hungarian) or in a foreign language (i.e. Turkish). The story was taken by the study of Buttelmann and his colleagues (2013), and we followed the same procedure with the exception that our procedure was carried out live. The story, translated to Hungarian and Turkish, were as following in respective languages, and in English:

Hungarian: *Szia Baba szia! Hogy vagy? Elmesélek neked egy történetet. Volt egyszer egy kacsa, aki egy tó környékén élt. Egy napon a kacsa útra kelt, hogy más állatokat találjon. Hosszú séta után a kacsa találkozott egy cicával. 'Szia cica!', mondta a kacsa. 'Milyen puha a bundád! Gyere, mutasd meg, hol laksz!' És mindketten elindultak a farmer házába.*

Turkish: *Merhaba bebek merhaba! Nasilsin? Simdi sana bir hikaye anlatacagim, bak soyle basliyor: golde yasayan bir ordek varmis. Bir gun, ordek baska hayvanlari* *bulmak icin bir yolculuga cikmis. Uzun bir maceradan sonra ordek, bir kediyle tanismis. 'Merhaba kedi!' demis ordek, 'tuylerin ne kadar da yumusak. Hadi gel, bana nerede yasadigini goster!' Sonra ikisi de ciftlige dogru yola cikmislar.*

English: *A duckling lived near a pond. One day, the duckling went on a journey to find other animals. After a long walk the duckling met a cat. 'Hello cat!' said the duckling, 'Your fur is really soft! Come on, show me where you live!' And they both went to the farmer's house."*

When the story-telling event was over the demonstrator reached under the table for the light box apparatus and put it on the table in front her, out of infants' reach. After placing her hands visibly on the table next to the two sides of the apparatus the demonstrator leaned towards the light box to get her forehead in contact with its surface for two seconds in order to turn on the light with her head. After bringing about the light-effect from the box she looked back at the infant, smiled and put her hands back on her lap. She then again visibly placed her hands on the table next to the apparatus, smiled at the infant, and leaned on the lightbox to demonstrate the head-touch action. There were three demonstrations in total. We followed the exact procedure of Kiraly et al.'s (2013) hands-free head-touch demonstration except that the demonstrators in our study engaged with the participants by using only non-verbal communicative signals (establishing eye contact) during the demonstration. The demonstration took approximately 20 seconds. When the demonstration was over, the demonstrator pushed the lightbox towards the infant without saying anything and left the room. Infants were allocated with 20 seconds of response period.

For the no demonstration condition, the experimenter escorted the dyad into the testing room, and she pushed the lightbox that was placed on the other end of the table to in front of the infant in the same manner as it was in the experimental conditions and then left the room. Parents were instructed not to intervene with their infants' behavior. Please see Figure 2.1 for the visual depiction of the experimental set-up for same- and foreign-language speaker conditions.

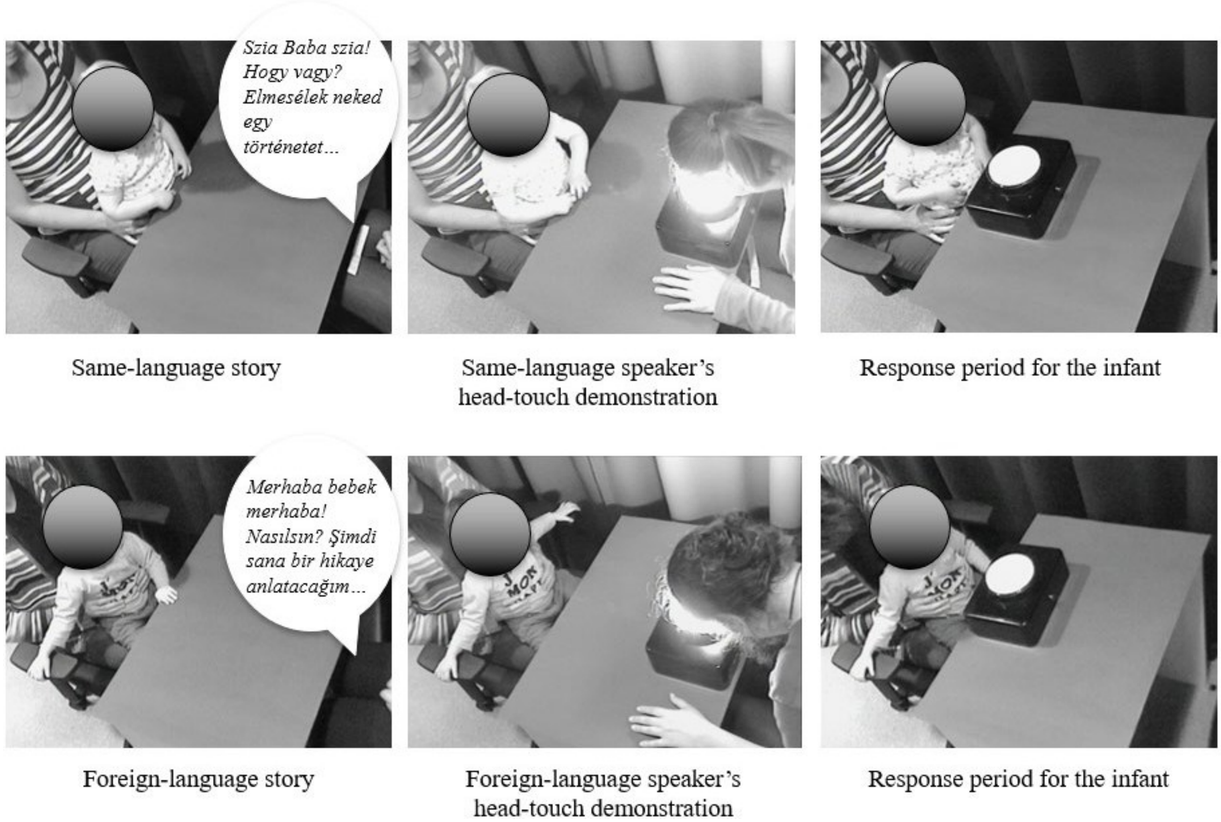

Figure 2.1. Experimental set-up for same and foreign language speaker conditions

2.6 Coding

We coded whether infants contacted the lightbox with their heads and their hands, and in which order. We additionally coded whether any actions infants performed were successful in bringing about the light effect. We had a 20-seconds-long response period following Meltzoff's (1988) and Kiraly et al.'s (2013) coding protocol. Response period started as soon as the baby touched any part of the apparatus (i.e. including the rectangular box the lamp was attached to). We also calculated the latency by registering how much time has passed from the moment the demonstrator put the apparatus in front of the infant until the infant's first touch of the apparatus.

A hand touch action was coded when there was a clear contact between the surface of the light box and the infant's hand. Exploratory hand behaviors, such as grasping or pulling the sides of the lamp in an attempt to remove it out of the apparatus, were not coded as hand-touch action. Note that such exploratory behaviors only took place in the no-demonstration condition. Please see Figure 2.2 depicting infants' different behaviors on the apparatus, along with an example of hand touch action on the surface of the light box. Following the standard set by Meltzoff (1988) and Kiraly et al. (2013), a head-touch action was coded either when there was a clear contact between the surface of the light box and the head or when there was a clear lifting or leaning action by the infant which resulted in decreasing the distance between the head and the light box to be 10 cm or less.

Figure 2.2. Hand-exploration taking place in the no-demonstration condition as seen in the first two pictures (a, b) along with an example of hand-touch action on the surface of the light box (c).

We further coded the frequency of hand and head touch actions. An action was considered to have ended when the hand (or the head) of the infant was retracted from the surface of the light box, and a second action began when she or he touched the surface again. Frequency coding was carried out independently of whether any of the actions (or actions attempts) caused the light box to light up or not.

31% of the data (n = 15 infants, 5 from each condition) were coded offline by a coder who was naïve to the hypothesis of the study and who was also blinded to which condition the infant was assigned by only watching the response period. Inter-rater agreement was excellent (Cohen's Kappa = 1) for the binary coding for head and hand touches, the order in which these were performed by the infant, and whether there was a light effect in the response period at least once. Inter-rater agreement was excellent for the frequency of head-touches ($r = 1$, $p < .001$) and very high for the frequency of hand-touches ($r = .94$, $p < .001$). Inter-rater agreement on latency measurement was also exceptionally high ($r = 1$, $p < .001$) with mean absolute difference between two codings being 113.73 ms ($SD = 184.2$ ms).

2.7 Results

We first analyzed whether there is any difference in the percentage of time infants attended to the story-telling events and to the demonstrations in experimental conditions, along with the latency in which infants acted on the box in three conditions.

2.7.1 Overt attention to story-telling and to demonstrations, and response latency

In order to make sure that infants equally attended to both the story-telling events that took place either in their own language or in a foreign language and the head-touch demonstrations we first measured infants' overt attention to story-telling events and demonstrations between experimental conditions. Percentage of time the infants attended to the demonstrator while telling the story was not statistically significantly different between groups ($t\,(29) = 1.68$, $p = .1$; Mann-Whitney $U = .81$, $p = .13$). Infants in both conditions looked at the demonstration for its entire duration too (medians in both conditions = 100%). Importantly the duration of the story telling events and of the demonstrations in the two conditions were also very similar in length despite live presentation between groups (for story telling events: $t\,(29) = 1.02$, $p = .31$; Mann-Whitney $U = .104$, $p = .54$; for demonstrations: $t\,(28) = .75$, $p = .46$; Mann-Whitney $U = .92$, $p = .4$).

A one-way ANOVA revealed a statistically significant difference in the latency with which infants acted on the light-box apparatus between three conditions of same-language speaker, foreign-language speaker and no demonstration ($F\,(2, 44) = 3.41$, $p = .04$, partial $\eta^2 =$

.13). A Tukey post hoc test showed that infants took more time to operate the light-box after the demonstration delivered by the foreign-language speaker (*M* = 10.51 secs, *SD* = 14.96) than infants in the no-demonstration condition (*M* =.93 secs, *SD* = 1.82; *p* = .04) but not than infants in the same-language speaker condition (*M* = 7.55 secs, *SD* = 10.54; *p* = .71). There were no differences in the latency with which infants acted on the lightbox between same-language speaker and no demonstration conditions (*p* = .19). Given the latency with which infants acted on the light box between two experimental conditions were not statistically different from each other, it appears that the foreign language spoken by the demonstrator did not lead the infants in the foreign-language speaker condition to take more time to act on the box in comparison with infants in the same-language condition.

2.7.2 Selective imitation of head touch actions on the lightbox

Despite attending equally well to the head-touch demonstrations delivered either by the same language speaker or a foreign language speaker, more infants in the same-language speaker condition performed head-touch behaviors. There were 13 infants (81%) in the same-language speaker condition and there were only four infants (25%) in the foreign speaker condition who acted on the lightbox with their heads at least once during 20-seconds coding window. All head-touch imitations were faithful to the demonstrations, no infant attempted to lift the lightbox towards their head. In the no-demonstration condition, despite not viewing a head-touch action on the lightbox, there were two infants (12.5%) who acted on the lightbox with their heads. To test whether the number of infants performing head-touch was affected by the condition, Fisher's exact tests were run with a Bonferroni correction for multiple comparisons (*p* = .016). There was a statistically significant difference between the experimental conditions (*p* = .004). Further, the number of infants performing the head-touch action was statistically significantly different between the same-language speaker and no demonstration conditions (p = .0002), but not between the foreign-language speaker and no demonstration conditions (p = .65)

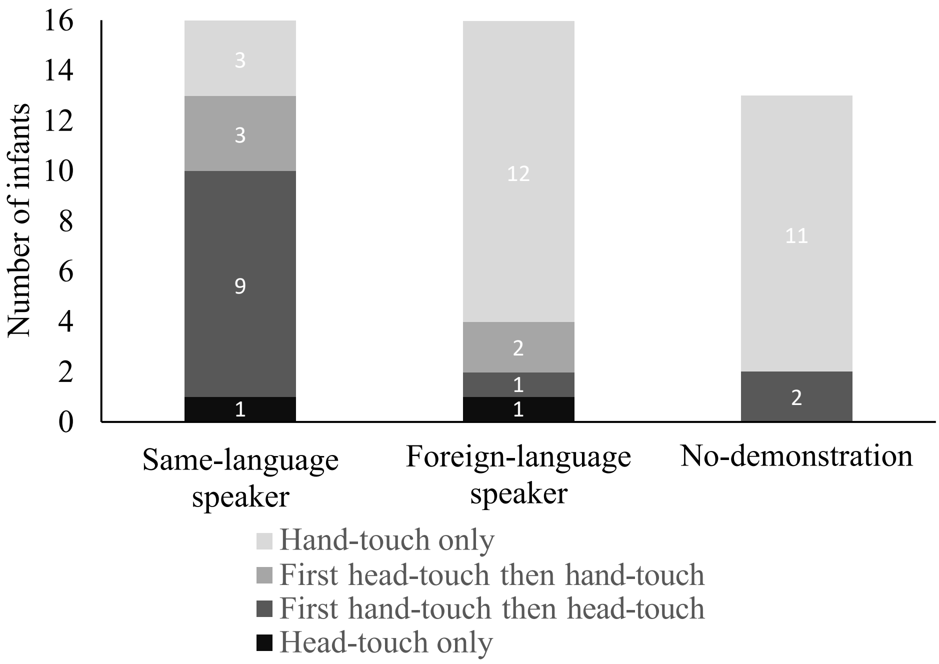

Figure 2.3. The number of infants who displayed only the head-touch action, first hand-touch and then head-touch action, first head-touch and then hand-touch action, and only the hand-touch action in three conditions. Note there were three infants in the no-demonstration condition who did not act on the lightbox neither with their hands or heads (but only engaged in exploratory behaviors on the box).

In addition to showing selective imitation of head-touch action, 15 infants out of 16 in both experimental groups and 13 infants in the no-demonstration condition acted on the lightbox with their hands at least once (see Figure 2.3). Yet we should note that not all actions infants performed were successful in bringing about the light-effect. Overall, in the same-language speaker condition there were 15 infants who successfully brought about the effect at least once, in the foreign-language condition there were 13 infants, and in the no-demonstration condition there were only eight infants doing so. Binary logistic regression indicated that the condition was a significant predictor for whether infants brought about a light effect or not ($\chi^2(2) = 8.88$, p= .01). The model explained 25.0% (Nagelkerke R^2) of the variance in bringing about the light effect correctly classified 75.0% of cases. This difference was followed up with one-tailed Fisher's exact tests given that we predicted lower rates of successful attainment of the effect in the no-demonstration condition in comparison to experimental conditions. There was a statistically significant difference in the number of infants bringing about the light effect between the same-language speaker condition and the no demonstration condition (p = .007, one-tailed), and a trend between the foreign-language speaker condition and the no demonstration condition (p = .067, one-tailed). There was no statistically significant difference between the two experimental conditions (p = .3, one-tailed).

2.7.3 Performing head touch action even after bringing about an effect with hand?

Critically to our hypothesis, we further investigated the number of infants who acted on the light box with their heads despite successfully bringing about the light effect with their hands before that. There were nine infants doing so in the same-language condition. In the foreign-language speaker condition if infants were successful to bring about the light effect with their hands first, none did perform the head-action later. Binary logistic regression indicated that the experimental condition was a significant predictor for whether infants performed a head touch after illuminating the box with their hands ($\chi^2(1) = 16.09$, $p < .001$). The model explained 57.0% (Nagelkerke R^2) of the variance in head-touch response after a successful hand action and correctly classified 78.1% of cases.

2.7.4 Frequency of hand and head touch behaviors on the lightbox

On average infants performed 6.69 (SD = 2.57) actions in the same-language speaker and 6.19 (SD = 4.08) actions in the foreign-language speaker, and 4.31(SD = 3.3) in the no-demonstration conditions. There were no differences between conditions in the number of actions infants performed both as revealed by parametric (F (2, 48) = 2.2, p = .12, partial η^2 = .09) and non-parametric tests ($\chi^2(2) = 4.43$, p > .05). Given that our frequency data for both measures were not normally distributed for any group, and also taking into account that there were more infants (n = 13) in the same-language speaker condition performing a head-touch action in comparison with other two conditions ($n_{foreign}$ = 4, $n_{no\text{-}demonstration}$ = 2), we calculated the ratio of head to hand actions for each infant in three conditions (for the raw mean frequency of head and hand actions please see Figure 1 on Appendix 1). Due to the one infant not performing any hand action on the lightbox in each experimental condition, the sample size for both the same-language and the foreign language condition was 15. Due to three infants not performing any hand touch action on the lightbox in the no demonstration condition the sample size for the ratio measure was 13 for this condition.

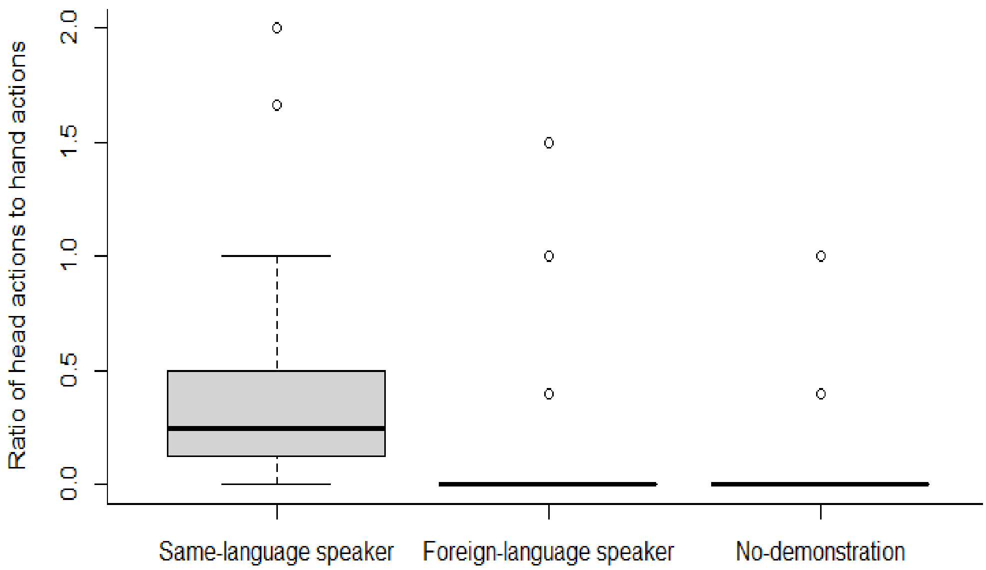

Figure 2.4 Boxplot for the ratio of head to hand actions by infants in the same-language speaker, foreign-language speaker, and no demonstration conditions

As Figure 2.4 shows mean ratio of head to hand actions was .47 (SD = .61) for the same language speaker, .19 (SD = .45) for the foreign-language speaker condition and .11 (SD = .29) for the no-demonstration condition. A Kruskal-Wallis test indicated a statistically difference in the means ranks of head to hand ratios between three conditions ($\chi^2(2) = 11.62$, p = .002). We followed up this difference with three non-parametric tests (significance level was Bonferroni corrected, p = .016). There was a statistically significant difference in the head to hand action ratios between same language and foreign language conditions (Mann-Whitney $U = 53.5$, $p = .008$), and between same language and no demonstration conditions (Mann-Whitney $U = 39$, $p = .003$). There was no difference in the head to hand action ratios between foreign language and no demonstration conditions (Mann-Whitney $U = 92$, $p = .82$ for foreign and no demonstration).

2.8 Discussion

The present chapter investigated selective imitation of sub-efficient action manners in 18-month-old infants. We hypothesized that an ostensive demonstration of a cognitively opaque action, such as leaning on a touch-sensitive light box to operate it with the forehead, will allow infants to encode the mean actions differently depending on the demonstrator's characteristics. Exploiting infants' early sensitivity to the informants speaking in the same-language as their own (Kinzler et al., 2007; Begus et al. 2016) we predicted that same-language would be a cue that infants can rely on while evaluating the relevance of the communicated action routine for their own social learning goals. Specifically, we hypothesized that 18-month-old infants will reenact sub-efficient actions that was demonstrated to them by a speaker of their own language. We also hypothesized that they will use an efficient alternative means in order to bring about the goal when the sub-efficient demonstration on how to illuminate the box was delivered by a foreigner.

Our results replicate Buttelmann and his colleagues' (2013) findings with a sample of older age group tested with a live paradigm. In line with our hypothesis, we found selective imitation of sub-efficient demonstration as a function of the language the demonstrators spoke. 81% of infants in the same-language speaker condition reenacted the head-touch action demonstrated to them in comparison to only 25% in the foreign-language speaker condition. Critically, as it was in the original study of Buttelmann et al. (2013), the language the demonstrators spoke was decoupled from their demonstration. Infants viewed the informant who told them a story in their own language or in a foreign language before the demonstration took place. Further, demonstrations were equally ostensive, hence communicative intentions of both models were readily available for infants to identify. The informative content, as both models displayed the same sub-efficient action on the same novel lightbox non-verbally, was also identical between the conditions. Regardless, infants took into account whether the demonstrators initially revealed themselves as speaking the same language as the infant or a foreign language and reenacted their respective demonstrations selectively even in the absence of those demonstrators.

We extended the original study on which our procedure was based (Buttelmann et al., 2013) by incorporating additional measures, such as using hand-actions and the relative frequency of head and hand actions. In the original study the authors reported that all infants in both same- and foreign- language groups acted on the box with their hands at least once but did not document when in the response period hand-actions were performed. If the infants consider the goal of the demonstration as achieving the light effect, there appears no reason for them to perform a head-action exactly as it was demonstrated to them, after they operate the lamp with their hands obtaining the effect. Instead, approximately half of the infants in the same-language speaker condition (n = 9) proceeded to light up the apparatus with their heads despite already having brought about the light effect with their hands before. In the foreign-language speaker condition, there was no infant showing this pattern. Only one infant operated the box with her head after performing a hand action in the foreign language speaker condition, and her first hand-action was not successful in illuminating the box. Our ratio analysis further revealed that infants in the same language condition were more likely to interact with the box using their heads in contrast with their hands, in comparison with the infants in the foreign language condition. This finding constitutes further evidence indicating that infants were not only inclined to selectively learn from the speakers of their own language, but they were also more likely to preserve the sub-efficient action manner that the same language demonstrator presented to them despite being able to operate the lamp with their hands.

These findings document that infants in the same-language speaker condition did not simply perseverate using the hand action strategy that enabled them to successfully to illuminate the lightbox in the first place. Rather the demonstration by the same language speaking model enabled infants to encode the precise manner in which the goal was achieved as revealed by the greater rates of high-fidelity imitation in the same language condition in contrast to the foreign language condition. However, we should note that two infants, in the no demonstration condition, despite not witnessing a head-touch action in the experiment, also performed a head action on the box. Given the flexible head-touch coding criteria we adopted from Meltzoff (1988) and Kiraly et al. (2013) that enabled us to also code infants' attempts to light up the box with their heads as long as the distance between the surface of the light box and their foreheads was 10 cms or less, and also considering the frequent occurrence of exploratory mouthing behaviors by infants around our targeted age, we attest that our liberal coding schema might have been the reason for the manifestation of head-touch action in the no demonstration group. Nevertheless, our findings on infants' selective imitation of the informants speaking the same language as the infants could be interpreted as even stronger evidence when one considers that the number of infants performing a head touch between the groups of foreign language and no demonstration conditions were similar to each other, as if the infants in the foreign language group had not seen an adult informant ostensively demonstrating the head touch action for them. Moreover, our ratio analysis also confirmed that infants in the foreign-language speaker condition did not interact with the light box differently than the infants in the no-demonstration condition. Infants in both conditions mostly acted on the box with their hands in comparisons with their heads. This was in stark contrast with infants in the same-language speaker condition.

Given the similar pattern of findings in the foreign-language speaker and no demonstration conditions, one could suggest that infants in the foreign-language speaker condition did not learn anything at all from a foreign language speaking demonstrator. However, while the rate at which infants successfully brought about the light effect was high for the infants in the same and foreign language speaker conditions (94% and 81%, respectively) infants in the no-demonstration condition was overall not very successful in bringing about the light effect. Only 50% of infants successfully operated the box in the no demonstration condition. This alone suggests that demonstration had a facilitatory effect in enabling infants to learn about object functions. The language of the demonstrations played a different role in learning though, as it was not a factor by itself simply influencing infant's learning about object functions as infants in both experimental conditions were able to illuminate the box regardless of the language of the demonstrations.

To sum up, our findings suggest that infants evaluated the demonstrators' epistemic competence and this evaluation influenced their decision as to whether they should incorporate the demonstrated actions in their cultural repertoire. When they perceived the model as an epistemically reliable source who could offer culturally shared knowledge that is relevant for them to acquire, they imitated the demonstrated sub-efficient action even in the absence of the demonstrator. It is likely that they encoded the head-action as a conventional culture-specific manner of how the lightbox ought to be illuminated. On the other hand, when the ostensively demonstrated action was rendered irrelevant for their cultural learning goals because it was presented by a foreign language speaker, they opted out from imitating the head-touch action (in the end it was a sub-optimal way of bringing about the goal from a teleological perspective), and relied on the efficient way they discovered themselves.

It is important to note that Howard and her colleagues (2015) could not find selective imitation for the sub-efficient actions delivered by an informant speaking the same language of the infants in a live demonstration with 19-month-olds. Critically, in their first experiment, they always had an experimenter speaking in the same language with the infants present during the imitation phase who encouraged them to act on the target apparatus independent of the condition they were allocated to. The presence of another adult, despite belonging to the same social group with the infants, could be the reason why infants in their study imitated the foreign-language speaker at similar rates with the same-language speaking informant. As covered exhaustively by previous research (Kiraly, 2009a; Over & Carpenter, 2012a; Kupán et al., 2017; Nielsen & Blank, 2011; Marsh, Ropar, & Hamilton, 2019), presence of an adult experimenter influences imitative behaviors of infants and children. Despite speaking a foreign language, the ostensive cues that the demonstrator exploited (e.g. making eye contact, infant directed intonation in speech) might had played a strong role in Howard et al. study as motivating infants' imitative behaviors to be in alignment with the foreign demonstrator's especially when there was social pressure to do so due to the presence of another adult during the reenactment phase. Importantly in our experiment there was no experimenter present in the response phase. Hence it is not likely that social-affiliative motives alone could be the explanatory factor justifying different patterns of imitative behaviors of infants we documented in the two experimental conditions.

Learning sub-efficient action manners selectively from same-language speaking informants despite being able to efficiently bring about the desired goal in the absence of the informants suggests that humans epistemically evaluate the informants based on the language they speak and guide their imitative behaviors accordingly from early on. They rely on self-discovered efficiency if they construe the demonstrated content as being irrelevant for their learning goals, and they incorporate ostensively demonstrated sub-efficient action manners into their knowledge repertoire if they interpret the demonstrated content as being relevant. It is an exciting avenue for future research to explore how exactly same-language vs foreign-language spoken prior to a demonstration leads to differential encoding of the same content as our findings cannot be explained by differences in overt attention allocated to different models. For now, we suggest that same-language spoken by the model might help infants to encode a novel arbitrary mean action as a cultural sub-goal given its sub-efficiency in relation to the end-state. Despite having a small sample size, this study provides additional evidence for cultural learning taking place early in ontogeny and proposes that this unique human tendency of learning sub-efficient action means from the same cultural group members is evolved to support the transmission of normative cultural knowledge shared within one's social group.

Chapter III. Infants' sensitivity to sub-efficient means actions shared within the language community

While even young infants show sensitivity to the efficiency of goal-directed actions and expect a rational (efficient) goal-approach from intentional agents (Gergely & Csibra, 2003), they also readily acquire sub-efficient actions, in which deviation from the efficiency is not justified by the physical environment, when these actions are demonstrated by social models (e.g. Meltzoff, 1988; Gergely, Bekkering, & Kiraly, 2002; Schwier, van Maanen, Carpenter, Tomasello, 2006). I argued that our uniquely human tendency to imitatively learn sub-efficient means actions from informants (e.g. as shown by "head-touch" and "over-imitation" studies, Meltzoff, 1988; Gergely et al., 2002; Kiraly, Csibra & Gergely, 2013; Gellen & Buttelmann, 2017, 2019; Nielsen & Blank, 2011; Horner & Whiten, 2005; Lyons et al., 2007, 2011) has evolved to support the fast intergenerational transmission of shared cultural knowledge in social groups (Gergely & Csibra, 2006; Legare & Nielsen, 2015). According to this view, infants and young children interpret such sub-efficient means normatively, as socially informative actions, characterizing how an action *ought to* be performed by the members of their social group. Therefore, infants can be expected to show high fidelity imitation when reenacting such actions, but only if they were demonstrated by agents who belong to their social group.

Chapter 2 presented supporting evidence for this hypothesis, indicating that infants selectively learn sub-efficient means from ostensive informants who speak the same language as their own social group. I argued that this can be taken as evidence that infants construe the sub-efficient manner in which the ostensively demonstrated means action is performed (activating a touch sensitive light-box by contacting it with head) as a relevant sub-goal (contact with head) that conveys culturally normative information about how the means action ought to be performed ("*this is how we do it*") (see also Kiraly, 2009b; Kiraly et al., 2013). If infants construe the ostensively demonstrated sub-efficient mean action as a sub-goal, that would represent the culture-specific manner in which the end state should be attained, then they should expect other agents who speak the same language as their own to attain the same end-state by performing the same sub-goal. However, we assumed that they would not generalize this expectation to foreign language-speakers. This chapter aims to empirically investigate this hypothesis, which, in more general terms, concerns infants' propensity to attribute a shared cultural knowledge repertoire to agents who belong to their own social group.

3.1 Introduction

Several different theories of goal understanding argue that goals are represented hierarchically, involving various sub-goals (Bekkering, Wohlschlaeger, & Gattis, 2000; Byrne & Russon, 1998; Gattis, 2002; Gleissner, Meltzoff, & Bekkering, 2000; Hamilton, 2009; Whiten, Flynn, Brown, & Lee, 2006). Such a hierarchical representation entails that an action can be described at different levels in relation to the goals that could be attached to it (e.g. from rotating the crank handle to opening the window, to getting fresh air), and the particular context in which it is performed. For example, the act of someone lighting candles affords a different goal interpretation, depending on whether the room is pitch black, or the room is already bright (Legare & Nielsen, 2015, on two action interpretation modes as instrumental and ritual). Empirical work has demonstrated that from an early age, infants can represent goals hierarchically, but as long as it is clear to them that the observed sequence of actions are linked to each other in a causally efficacious way, such that one action is performed in order to "*establish necessary conditions*" for the second action (i.e. causally enabling relations, Kiraly, 2009b).

For example, in a looking time paradigm, Sommerville & Woodward (2005b) habituated 12-month-old infants to a scene in which an agent pulled a cloth supporting an object that was placed on the far end of it, and then grasped the object. Infants then saw test events where this time the agent acted only on the cloth. In one test event the agent was acting on the cloth supporting the same object she grasped during the habituation, in the other test event the agent was acting on the cloth supporting a different object. Despite not viewing the grasping action in any of these test events, infants looked longer to the second test event, showing surprise upon seeing the agent acting on the cloth that supported a different object than the one they saw the agent repeatedly grasping during the habituation phase. This study shows that infants represented the pulling-the-cloth event as a sub-goal to bring about the hierarchically higher-level goal of obtaining the object. Critically, infants failed to do so when the causal link between cloth-pulling and grasping-the-object was violated, when the object did not rest on the cloth.

Infants' ability to recognize the actions as sub-goals in action hierarchy seems to be influenced by whether actions relate to the end state in the most causally efficacious way available. This finding was also replicated in behavioral paradigms showing young children's imitation of a demonstrated action sequence depends on whether or not means actions efficiently relate to the end state (Bekkering et al., 2000; Carpenter, Call, & Tomasello, 2005). For example, when 12- and 18-month-old infants viewed an agent placing a toy mouse into a toy house by making the mouse either to slide or hop to its final location, infants in both age groups imitated only the end goal of the action sequence, ignoring the manner in which the mouse was moved (Carpenter, Call, & Tomasello, 2005). Similarly, Bekkering and his colleagues in several studies (Bekkering, Wohlschlaeger, & Gattis, 2000; Gattis, Bekkering, & Wohlschlaeger, 2002; for a review, Wohlschlaeger, Gattis, & Bekkering, 2003), showed that young children omit reproducing the exact style with which object (or body-part) directed actions were demonstrated to them. In a "do as I do" imitation paradigm, when 3- to 6-year-old children were presented with a model performing ipsilateral and contralateral manual actions to reach a goal state (the model touched her right ear with her right hand or touched her left ear with her right hand) children were more likely to reenact the model's contralateral action with an ipsilateral action, while they all touched the correct ear. Thus, they emulated the goal, rather than imitating exactly the target action leading to it. The same response pattern was found when the gestures were directed to dots on a table (Bekkering et al., 2000). Importantly, in these studies the exact manner in which the demonstrated actions were performed (e.g. hopping or contralateral reaching) did not connect to the goal states (e.g. going into the house or touching the ear or dots on the table) in the most efficient way possible. Hence children responded by reproducing the final outcome of the action sequence, at the highest level of the goal hierarchy, without necessarily reenacting exactly the actions that formed the lower level sub-goals in the hierarchy.

However, several studies also showed that when the number of goals in the hierarchy are reduced infants can encode the goals lower in the hierarchy as well, even when the action appears arbitrary or does not causally connect to the end state efficiently. For example, when the toy house was removed from the demonstration altogether, Carpenter and her colleagues (2005) found that infants this time imitated the manner in which the toy mouse was moved across the mat. Similarly, when there were no dots on the table anymore (Bekkering et al., 2000) or when the manual actions terminated near the ear, rather than by touching the ear (Gleissner, Meltzoff, & Bekkering, 2000) young children construed the action itself as the goal. In a similar study, when children observed the model only touching her right ear with ipsi- and contra-lateral gestures, instead of both of her ears alternately, children this time faithfully reproduced the actual target action (Bekkering et al., 2000). As these studies show, removing the final goal or having the same end state achieved through multiple actions induced young children to encode the actions themselves as the goals of the agent or the sub-goals of the agent in a goal hierarchy.

Infants' reliance on certain social cues from early on in ontogeny also enables them to attend to goals lower in the hierarchy. For example, when the agent demonstrated a peculiar sub-efficient action, such as leaning over on a touch sensitive light box to bring her head into contact with it to make the box light up, infants were more likely to reenact this sub-efficient means when the action was demonstrated in an ostensive communicative context (Kiraly et al., 2013). This was the case even though the causal enabling relation between the sub-goal (i.e. make contact with the head) and the goal at the higher level of the hierarchy (i.e. illuminating the box) was cognitively opaque to the infants, as the demonstrator's action choice was not the most efficient means available to her. According to Natural Pedagogy theory (Csibra & Gergely, 2009), ostensive communicative signals (such as eye contact and infant directed speech) lead to differential encoding of communicatively presented information, by biasing the recipient to interpret the content of the demonstrated information as relevant and that should be learned and retained. This species-unique sensitivity to such ostensive cues is argued to be an evolved mechanism for relevance-guided cultural learning, enabling efficient intergenerational transmission of cognitively opaque forms of cultural skills and knowledge (Csibra & Gergely, 2011). According to this theory, ostensive signals provides a context which indicates to the infants that the action sequence demonstrated and the manner of means action through which the final goal in the hierarchy is achieved represent culturally relevant information for the child to acquire (Kiraly, 2009b; Kiraly et al., 2013; Brugger et al., 2007).

Further evidence on the role of ostensive signals in enabling infants to attend to the goals lower in the hierarchy comes from Southgate, Chevallier and Csibra (2009). Employing the same paradigm as Carpenter et al. (2005) they presented to 18-month-old infants a toy animal either hopping or sliding into a toy house. However, critically, for one group of infants the demonstrated outcome of the action (i.e. animal going into the house) was not new information as it had been already communicated ostensively before the action demonstration.

Another group of infants did not receive such prior information about where the toy animal would go, and another group received the same prior information about the outcome of the action before the demonstration, but this information was delivered non-ostensively. Only the infants in the ostensive-prior information group imitated the manner in which the toy animal was moved to the house, constructing the manner of movement as the relevant sub-goal to be learned. When all the action elements (manner and outcome) were equally novel to the infants or when the manner information was not accompanied by ostensive signals, infants imitated the demonstration at the hierarchically highest level, in terms of its outcome.

These studies show that ostensive signals enable infants to interpret the demonstrator's communicative intent as being about the novel components of the action demonstration worthy of acquiring. Hence, when the action manner is novel and is accompanied by ostensive signals, the manner becomes the relevant piece of information construed as a sub-goal and is reproduced with high fidelity (Southgate et al., 2009; see also infants' performance in the second test trials in Schwier et al., 2006). Although the reason why the demonstrator chose that particular sub-goal might not be transparent to the observer, especially given that she could have more easily performed an efficient action instead of its sub-efficient manner variant, young children seem to understand that there must be a relevant reason for the actor's sub-efficient action display (Schwier et al. 2006, Gergely & Csibra, 2006; Williamson & Markman, 2006)[3]. It appears that when the use of ostensive signals indicates to the infants that the actions leading to the goal are relevant, they show readiness to adopt them despite their sub-efficiency and reproduce them exactly. This ability to recognize the relevance of a cognitively opaque action sequence that is ostensively demonstrated is in fact very important as it allows naïve

[3] As adults we also intuitively rely on others' cognitively opaque means actions in unfamiliar situations, for example closely copying the other person at a formal dinner table in choosing which utensils to use for which dish or in deciding how to consume an unfamiliar dish.

juveniles to learn about the social conventions and cultural practices of their cultural community.

3.1.1 Why prioritize means actions in action hierarchies?

According to Watson-Jones & Legare (2016) recognizing cognitively opaque actions, as, for example, in the case of rituals, might solve problems associated with living in groups, e.g. identifying other agents as group members. Being able to predict someone else's social group membership has utmost importance given the benefits of forming and maintaining coalitions, coordinating group members for group goals, and ensuring commitment from the same-group members (Chudek & Henrich, 2011; Watson-Jones & Legare, 2016). However, to date there has been only one study investigating the role of causally opaque means actions in infants' inferences about third party social relations. Liberman, Kinzler, and Woodward (2018) had 16-month-olds view video segments showing two individuals adopting different sub-efficient means in their goal attainment (one individual used her forehead to activate a touch sensitive light-box apparatus to emit the light effect while the other pressed her elbow on the surface of the box to do the same). After these familiarization scenes, in one test phase infants watched these two actors signaling their positive affiliation towards each other, and in another test scene they watched the two actors as they signal their disengagement from one another. Infants looked longer during the affiliation test phase in comparison to the disengagement test phase; showing their surprise to see the two individuals, who acted on the same object by different causally opaque means, to *affiliate* with each other.

In a follow-up experiment, Liberman and colleagues (2018) manipulated whether one of the two agents the infants observed acted under a physical constraint or not. In the "blanket" condition, one of the agents wrapped a blanket around her shoulders, causing her hands to be occupied, similarly to the original study of Gergely et al. (2002). Given her physical constraints, the agent illuminated the lightbox by a head touch. The other agent the infants viewed used her hand instead to light up the box as she did not have the same constraints as the other agent. In the "no-blanket" condition, infants viewed one of the agents purposefully performing the head-touch action, without any constraints, whereas they saw the other agent activating the box with her hand. Upon watching the same affiliation and social disengagement test videos as in Experiment 1, infants in the no-blanket condition looked longer when the agents, who both purposefully used different means to activate the box, affiliated with each other. In other words, when the agents freely chose different means to activate the box infants did not expect them to affiliate but expected to see them showing disengagement from one another. Infants in the blanket condition looked longer when the agents displayed their social disengagement in contrast to when they displayed their engagement with one another. In other words, when both agents performed the same goal but just had to use different means while doing so due to presence or absence of the physical constraints, infants were surprised to see their social disengagement in the later test trials.

This study further indicates that infants could construe the goals of others differently depending on contextual cues. In one case, when there was a physical constraint compelling one agent to adopt an (otherwise) sub-efficient means, but when a similar constrain was not there for another agent, 16-month-old infants construed both agent's goals as illuminating the light box. Whereas when a sub-efficient action manner was performed by one agent purposefully and freely, and an efficient manner was used by the other agent, they construed the first agent's goal behind the head action as socially motivated. Since the actor intentionally chose to adopt the sub-efficient means action, it was clear to the infants that she did not only act to achieve her goal of bringing about the light effect, but she also had the sub-goal of doing so in a specific way.

However, the finding that infants can infer the social relationship between two agents based on the action manners they purposefully adopted does not necessarily imply that the agents can be construed as belonging to the same social group. It is also unclear, as the authors themselves pointed out (Liberman et al., 2018), whether infants expected the two agents not to affiliate with each other because they were simply not imitating each other (in the no-blanket condition), or because purposeful sub-efficient manner display could signal socially rich information about the conventional way that the activity should be performed, which in turn would indicate one agent's knowledge about that convention, and another's lack of knowledge in the same domain.

Here we aimed to provide a more direct empirical test of the latter hypothesis. Based on previous research demonstrating that one reliable cue that infants from an early age attend to when forming social category representations is spoken language (Liberman, Woodward, & Kinzler, 2017, Soley & Aldan, 2020), in the present research we used same versus foreign language as a cue to induce the representation of social group membership in infants. Social group membership based on language could mark the boundaries of shared cultural knowledge, indicating that speakers of the same language would likely to share conventional practices of their own cultural communities while the speakers of another language would likely to share their own (Soley & Aldan, 2020; Olah, Elekes, & Kiraly, 2019). Hence in this chapter our aim was to probe infants' understanding that shared cultural knowledge is delineated by linguistic social group membership. If infants are sensitive to the cue of shared language as an indicator of shared cultural practices among agents, then they may be able to draw inferences about the conventional action manners the agents from the same linguistic social group would adopt.

3.2 Present research

Sub-efficient means actions (e.g. as in "head-touch" studies where a social model lights up a touch-sensitive box by leaning on it to bring her forehead in contact with its surface, Meltzoff, 1988; Gergely, Bekkering, & Kiraly, 2002; Kiraly, Csibra & Gergely, 2013) are good candidates for conventional practices. Given their cognitive opaqueness (i.e. they are not the most efficient means to fulfill the instrumental goal, hence the reason behind its performance is not transparent) (Gergely & Csibra, 2006), observers cannot rely on efficiency considerations to interpret such manner actions, but only on the basis of relevant social and contextual cues can they make sense of the sub-efficient actions they observe. Considering also that sub-efficient actions require agents performing them to extend extra effort in doing so, sub-efficient means actions, just like rituals, could signal commitment, i.e. indicating the agents' adherence to somewhat costly practices of their community (for a review see Watson-Jones & Legare, 2016). Given also their arbitrary features, how these action routines could have been acquired and retained to be also purposefully transmitted to others is not clear. After all, the likelihood of an unfamiliar agent discovering a seemingly arbitrary version of the means action to achieve the goal (while more efficient alternative to achieve the same goal could have been easily discovered) is very low, unless the agent acquired it from a knowledgeable cultural member who had already been familiar with the culture-specific skill variant that she herself is likely to have acquired from an ingroup source in the first place (rather than having discovered it independently).

Based on these assumptions, here we hypothesize that ostensive sub-efficient means demonstration on a novel apparatus could give a rise to encoding the manner as a cultural sub-goal by infants. We hypothesized that an ostensive demonstration of a sub-efficient action manner can induce an interpretive bias in young infants sanctioning the inference that the manner demonstration represents culturally shared generic information about how the apparatus ought to be activated. If so, infants could expect that the demonstrated sub-efficient head action would convey information that is generalizable beyond the episodic situation and the specific agent. Thus, we predicted that infants would encode the particular sub-efficient manner (i.e. an informant's leaning-over-the-light-box to illuminate it by contact with the head) that they have observed in an ostensive context as a sub-goal (i.e. head-contact) that provides normative information as to how the means action ought to be performed (i.e. should be illuminated with head). Consequently, infants would expect this sub-goal to be shared cultural knowledge available to all members of the given language community (e.g. Hungarian speaking agents) but not to the members of a different language community (e.g. Chinese speaking agents).

3.3 Experiment 1

Fourteen-month-old infants were tested in a looking time paradigm. Infants first viewed an ostensive novel sub-efficient action manner demonstration by an agent speaking in their own language (i.e. Hungarian) on a novel apparatus. This demonstration presented the informant leaning over the touch sensitive box to get her forehead in contact with its surface in order to bring about the light and sound effect emanating from the box. This episode intended to serve as a teaching event for the infants, aiming to induce an interpretation that this is how one should activate the apparatus.

Then infants watched two new agents entering the scene as they spoke with each other either in the same language as the informant (speaking in Hungarian: congruent language condition) or in a different language, which was unfamiliar and foreign to infants (speaking in Chinese: incongruent language condition) using a between-subjects design. Infants later viewed these two agents acting on the box by adopting different novel sub-efficient means. One agent used her elbow to bring about the effect (elbow action) and the other agent lifted the box emulating the head-touch action of the informant (lift-to-head action) (within-subjects). Both actions were carried out non-ostensively.

In earlier imitation studies using the head-touch paradigm it was reported that when the light-box was movable and small enough to be lifted by the infants, 14-month-olds tended to *lift the box to their heads by their hands* (rather than leaning over to touch the box by their head, as was done by the demonstrator) (Chen, Kiraly & Gergely, 2012). Based on this finding (which indicates an emulative response to reproduce the demonstrated sub-goal of "making contact between head and box"), we assumed that infants would interpret the lift-to-head action performed by the novel agent as being an equivalent emulative version of the initially demonstrated leaning-over-head-action that had been observed by them in an ostensive communicative context. Using a novel emulative variant of the demonstrated head-contact sub-goal during test also allowed us to test our hypothesis in a way that kept both test events perceptually novel. In contrast, we assumed that infants would represent the elbow action as a distinctively different (and non-equivalent) realization of the demonstrated sub-goal. Based on this argument, we predicted that infants would look longer (showing violation-of-expectation) at the completely novel means action (the elbow-action) performed by the agent in contrast to the equally novel, but emulative version (i.e. the lift to head by hands action) of the demonstrated sub-goal in the congruent language condition (where the three agents spoke the same language).

On the other hand, for the incongruent language condition we predicted infants to be surprised upon seeing the foreign language agent performing the more similar (emulative) version of the mean action that had been demonstrated by the informant using a different language than the test agents. Specifically, we expected 14-month-olds to look longer to the lift-to-head action in comparison to the elbow-action in the incongruent language condition. We argued that if infants are sensitive to the fact that the specific sub-efficient manner of means action demonstrated is likely to be only accessible (as shared knowledge) to those agents who have been socialized into the culture-specific practices of the same social group (as indicated by their shared language), they should find it unlikely that members of a completely different social group (speaking a different language) would also adopt the similar sub-efficient action manner.

3.3.1 Participants

Participants were 32 14-month-old infants (10 females). Mean age of the participants was 437.44 days (14 months 10 days) (range: 406-455 days; 13;10 to 14;29, *SD* = 9.7 days). An additional 20 infants were tested but excluded from the final sample due to experimenter error (*N* = 3), mother interference (*N* = 3) or fussiness (*N* = 5). We further excluded nine infants because they were looking at the test events less than 2 seconds in total. All infants were monolingual.

Participants were recruited from a database of parents who volunteered to participate in developmental studies. Each parent gave his or her written informed consent for the study, and the procedure was approved by the United Ethical Review Committee for Research in Psychology, Hungary, and conducted in accordance with the Declaration of Helsinki. Infants were given a toy or other small gift in the end of the study independent of their performance.

3.3.2 Design

Infants were randomly assigned to two between-subject conditions: congruent language (*N* = 16 [4 females], with the mean age of 437.37 days with a *SD* of 7.81 days) and incongruent language (*N* = 16 [6 females], with the mean age of 437.5 days with a *SD* of 11.54 days).

3.3.3 Procedure and stimuli

Infants were tested in a dimly lighted soundproof room. They sat on their caregiver's lap 100 cm away from a 40-inch plasma screen where the stimuli were presented via PsyScope X B77 software. A hidden camera mounted below the screen recorded infants at 25 frames per second. The caregivers were instructed to keep their eyes closed and to remain silent during the whole procedure. The stimuli were short movies edited in FinalCut Express 4.0 software. Infants watched three short familiarization movies followed by two test events. An attention-grabbing animation (black and white checkerboard moving in the center of the screen while emitting a ringing sound) was played before each movie. During the test events, the experimenter sitting at the control room assessed infants' looking behaviors toward the screen through the monitor and made a continuous mouse press when the infant was looking at the screen and released her press when the infant looked away. If the duration of the look-away was more than 2 seconds, the software continued presenting the next text event (please see coding section).

In the first familiarization scene infants saw a female adult sitting in the middle of a table with a touch-sensitive box in front of her, which beeped and lighted up upon being touched on the surface. The agent first greeted the infant in the infant's own language by looking directly at the camera, smiling and saying (in Hungarian) "Hello baby, hello. Now I am going to show you something. Look!" The ostensive informant then demonstrated a head-touch action on the touch-sensitive box by first putting her hands on the table and then leaning towards the box to get her forehead in contact with its surface to bring about the effect. The informant, after saying "Hello baby, hello. I am going to show you again. Look!" repeated this head-touch action demonstration.

Following this initial ostensive head-touch action demonstration (with the total running time of 18.32 seconds) two new agents entered the scene as one of them handed a paper tissue to the other while they were interacting with each other either in the same language the informant spoke, which was also the native language of the infants (i.e. Hungarian: congruent language condition) or in a different language, which was a foreign language for the infants (i.e. Chinese: incongruent language condition). The ostensive informant attended this interaction by looking back and forth between two agents. Decibel levels of these two spoken interactions and of the ostensive head-touch demonstration were matched to each other in Praat software. The interaction scene took 12 seconds and their verbal exchange (either in Hungarian [for the congruent language condition], and in Chinese [for the incongruent language condition]) was as follows:

Actor$_1$: [coughs] do you have a napkin?
Actor$_2$: yes, [reaching for her pocket] it should be here, [hands the napkin over]. Here you go.
Actor$_1$: thank you.
Actor$_2$: you’re welcome.

In Hungarian:
Actor$_1$: (köhög) van egy zsebkendőd?
Actor$_2$: igen, (belenyúl a zsebébe), itt kell lennie (átadja a zsepit)
Actor$_1$: köszönöm.
Actor$_2$: szívesen.

In Chinese:
Actor$_1$: (咳嗽) 你有餐巾纸吗? (用来擦鼻子或鼻涕)
Actor$_2$: 有 (手伸向裤子的口袋，去拿餐巾纸), 应该在这儿。给你
(将餐巾纸递向对方)
Actor$_1$: 谢谢！
Actor$_2$: 不客气！

The interaction scene was followed by two action events. In both action events the agents who had interacted with each other displayed two different novel sub-efficient means on the touch-sensitive box, respectively. One agent used her elbow to touch the lamp to bring about the light and sound effect ("elbow action"), while the other agent lifted the box by hand to her head bringing it into contact with her forehead to activate the box ("lift-to-head action") (see Figure 3.1 for depiction of the video stimuli). The lift-to-head action was similar (or equivalent) to the ostensive informant's head action in so far as they achieved the same sub-goal state ("contact between head and lamp") as the head-touch action originally demonstrated. In other words: both alternative means actions induced establishment of contact between the surface of the box and the forehead, but the actual action performed by the agent during test was a non-identical emulative version of realizing the same sub-goal state as it involved lifting the box by hand to her head to establish contact between head and lamp, instead of being an exact imitative copy of the originally demonstrated head-action (leaning forward to touch the lamp with the head). The alternative "elbow action" produced by the other agent involved the agent using her elbow to touch the surface of the box. For both action events, the body part used stayed in contact with the surface of the box which emitted a light and the sound effect for two seconds. The duration of these two action events was 11.08 seconds each. Both test events started as the agent raised the box from her lap and put it on to the table and ended when she placed her hands down on the two sides of the box after having performed the particular sub-goal action. Each agent performed their respective version of the sub-goal action twice and did so non-ostensively: neither agent looked at the camera, to the ostensive informant or to each other while acting on the box. The ostensive demonstrator watched each of the two alternative action events attentively. During the test trials the final still frame of each action event stayed on the screen for a total of 60 seconds or until the infants looked away from the screen for two consecutive seconds. Which actor displayed which action, and which action was carried out first were counterbalanced across participants, yielding four different counterbalancing conditions.

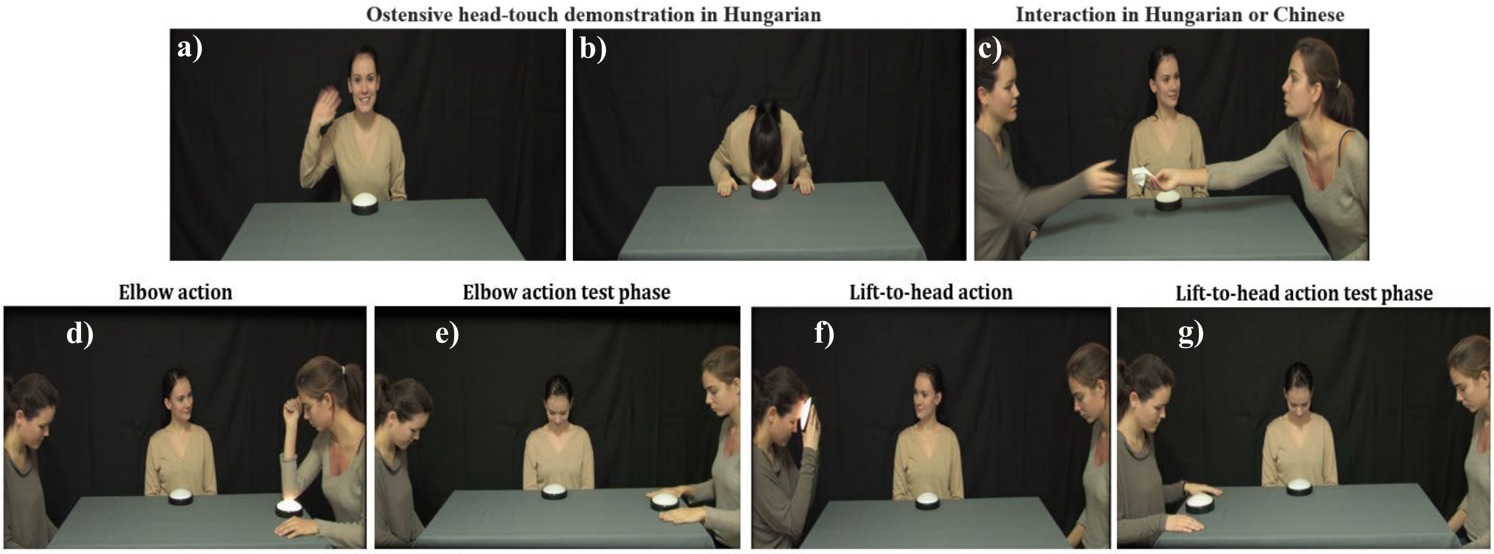

Figure 3.1. Depiction of the stimuli, showing the ostensive informant a) greeting the infant, b) displaying the head-touch action on the box, c) the two agents interacting with each other, d) the elbow action, e) the final still frame of the elbow action test event, f) the lift-to-head action, g) the final still frame of the lift-to-head action test event.

3.3.4 Coding

We coded infants' looking behaviors toward the screen frame-by-frame offline. Our coding criteria required infants to look at the screen at least for two consecutive seconds from the start of our looking time measure until they look away from the screen for two consecutive seconds or until 60 seconds passed. Infants had to observe at least one ostensive head-touch action demonstration fully, one elbow-action and one lift-to-head action to be included in the final analysis.

28% of the looking time data (nine infants) were coded by a naïve coder, blind to the hypotheses of the study. Inter-coder reliability was high (r = .99 for the two test actions collapsed, mean absolute difference between the codings for the test actions was 224.44 ms, SD = 559.68 ms).

3.3.5 Results

All 32 infants observed the ostensive head-touch demonstration, the interaction episode, and at least one action event for each sub-efficient mean display. They looked at the ostensive head-touch demonstration for 96.75% of its duration (*M* = 18.3 seconds, *SD* = .47 seconds), and at the interaction for 98.93% of its duration (*M* = 11.87 s, *SD* = .33 s). The language of the interaction did not influence infants' looking times at the interaction scene (t (29) = .26, p = .79, Cohen's d = .09; Mann-Whitney U = 116, p = .89). (Note that there was one infant whose video recording was missing for the interaction movie in the incongruent language condition.) Additionally, infants looked at the first action event for 95.5% of its duration (*M* = 10.58 s, *SD* = .85 s), and at the second action event 97.59% of its duration (*M*= 10.81 s, *SD* = .72 s). Sixteen infants saw the lift-to-head action first, while the remaining participants saw the elbow-action first. Which action was seen first did not affect infants' looking times at the two action events as the interaction of order of presentation and the action type presented indicated no significant effect (F (1, 30) = 2.47, p = .13, partial η^2 = .08). Fifteen infants saw agent A acting first while the remaining infants saw Agent B first. Which agent was seen first did not influence differently infants' looking times at the two action events (F (1, 30) = 2.53, p = .12, partial η^2 = .08). Infants looked at the lift-to-head action 95.54% of its duration (*M* = 10.86 s, *SD* = .57 s), and they looked at the elbow action 95.1% of its duration (*M* = 10.53 s, *SD* = .94 s).

For test trials we log-transformed the raw looking time durations given that the data were not normally distributed (Csibra, Hernik, Mascaro, Tatone, & Lengyel, 2016), and used an alpha value of .05 for all statistical analyses reported below. Data were collapsed across participants' gender as preliminary analyses showed no significant interaction of this factor with test trials (F (1, 30) = .15, p = .7, partial η^2 = .005). There were also no significant interactions between which action was first seen and the test trials (F (1, 30) = 1.45, p = .24, partial η^2 = .05) and between which agent acted first and the test trials (F (1, 30) = .61, p = .43, partial η^2 = .02).

An analysis of variance on the log-transformed looking times during test trials with the interaction language as a between subjects factor (congruent and incongruent language) and the action type (elbow-action and lift-to-head action) revealed a marginally significant effect for action type (F (1,30) = 4.17, p = .05, partial η^2 = .12), but no significant interaction (F (1,30) = .23, p = .63, partial η^2 = .008). Infants across the two groups looked longer at the lift-to-head action (M = 19.78 s, SD = 15.07 s) than at the elbow action (M = 15.29 s, SD = 14.06 s).

Planned comparisons revealed that infants in the congruent language condition looked at the lift-to-head-action significantly longer than at the elbow-action (t (15) = 2.23, p = .04, Cohen's d = .57). The difference between the looking times for the two different actions was not statistically significant in the incongruent language condition (t (15) = .95, p = .36, Cohen's d = .25). Figure 3.2a shows the pattern of looking times at the two types of test actions (reported in seconds) while Figure 3.2b illustrates the same contrast for log-transformed looking times.

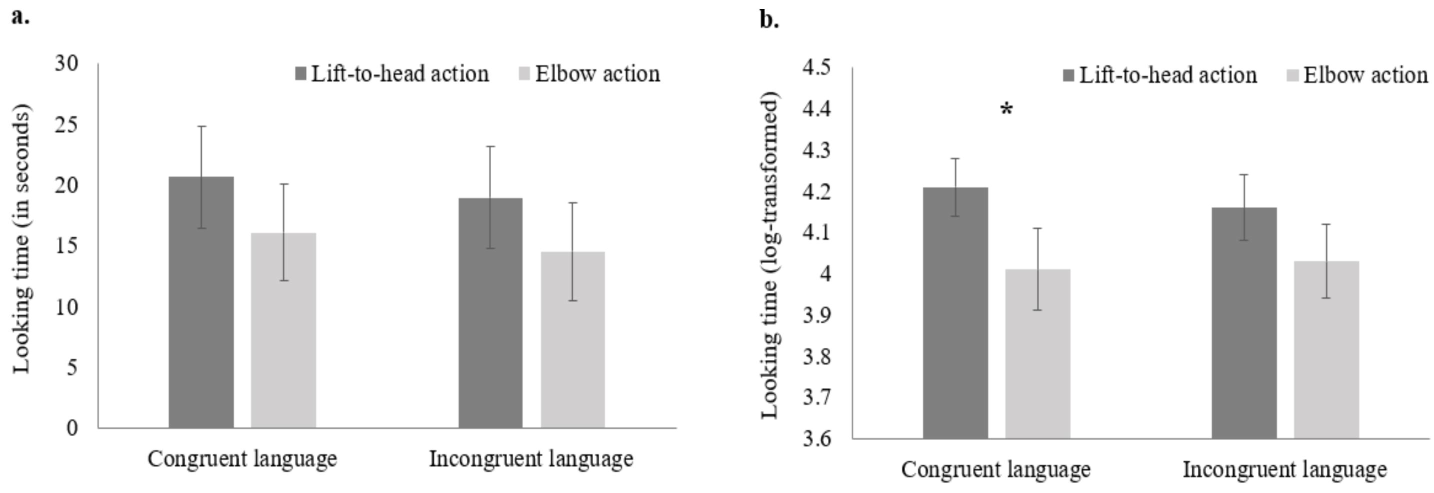

Figure 3.2a. Mean looking times in seconds at the two test trials for infants in both conditions in Experiment 1. **3.2b.** Mean log-transformed looking times at the two test trials for infants in both conditions in Experiment 1. Error bars show ±1 SE. Asterisk marks a statistically significant difference.

3.3.6 Discussion

Fourteen-month-old infants in both language groups looked longer at the lift-to-head action than at the elbow-action. The finding of a statistically significant pattern of longer looking at the lift-to-head action in the congruent language condition was not in line with our prediction. Below we discuss possible reasons for this pattern of findings.

One possibility is that our familiarization events were too complex to process for 14-month-olds. In contrast to the paradigm used by Liberman and her colleagues (2018) in our paradigm there were three different agents, doing three different actions. Additionally, infants in the congruent language condition observed two of these agents verbally interacting with each other in the same language that was spoken by the initial ostensive demonstrator while in the incongruent language condition they observed the two agents speaking an unfamiliar different language. Furthermore, our familiarization events for the ostensive teaching episode as well as our test events were also rather short in so far as the displayed actions were presented only twice (rather than eight times as in Liberman et al., 2018). Given the complexity of the sequence of actions we presented to 14-month-olds, it may be possible that infants' tendency to look at the "lift-to-head action" longer in our paradigm reflected their preference to look at the event they found familiar.

Indeed, previous research documented that infants prefer to attend to events that provide them with optimal level of predictability (Roder, Bushnell, & Sasseville, 2000; Kidd, Piantadosi, & Aslin, 2012; Kidd & Hayden, 2015). Scenes with very predictable elements (that include "very low information content") or scenes that are highly unpredictable (including "high information content" and as such become cognitively overly demanding for infants) do not represent optimal stimulus conditions for infants in their pursuit of information gain (Kidd & Hayden, 2015). Relatedly, Powell & Spelke (2018a) argued that when infants have fully encoded a scene, looking at it longer does not help them learn any new information; hence they might opt to look longer at a novel scene instead. However, when infants have not encoded a scene fully, and so they yet have more to learn about it, they may continue to look at it longer (despite its relative familiarity), than at a novel scene.

Given these considerations, it is possible that 14-month-old infants in this experiment looked at the “lift-to-head action” longer because our paradigm has failed to provide infants with an event sequence that could be encoded fully at that age. It is a possibility that their looking time behaviors reflected the processing demands associated with complex stimuli they were presented with, leading them to look longer to the event they found familiar based on the initial familiarization. If this were the case, then it may be possible that older infants with more processing capacity would show a different pattern of looking times than the 14-month-olds, and would look longer at the less expectable elbow action, a result that would be in line with our predictions. To test for this possibility, in Experiment 2, we increased the age of our sample. Also, this time we tested the older infants only in the congruent language condition.

3.4 Experiment 2

Experiment 2 used the exact same stimuli as Experiment 1. Participants were 18-month-olds and were tested in the congruent language condition only.

3.4.1 Participants

Participants were 16 infants (9 females). Mean age of the participants was 573.06 days (18 months 25 days) (range: 553-612 days; 18 to 20 months, *SD* = 21.53 days). An additional 15 infants were tested but excluded from the final sample due to being already tested with the same paradigm when they were 14 months of age (*N* = 2), experimenter error (*N* = 2), mother interference ($N = 4$) or fussiness ($N = 6$). We further excluded one infant because she was looking at the test events less than 2 seconds in total.

3.4.2 Procedure and stimuli

Procedure and stimuli were the same as for the Experiment 1 except that infants in Experiment 2 observed the three agents all speaking the same familiar language as their own (as in the congruent language condition of Experiment 1).

3.4.3 Coding

We coded infants' looking behaviors toward the screen frame-by-frame offline. Our coding criteria was the same as it was in Experiment 1. 25% of the looking time data (four infants) were coded by a coder naïve to the hypothesis of the study. Inter-coder reliability was high for action test trials ($r = .99$ for two test actions collapsed, mean absolute difference between the codings for the test actions was 31.11 ms, $SD = 26.67$ ms).

3.4.4 Results

All infants viewed the ostensive head-touch demonstration, interaction, and at least one action event for each sub-efficient mean display. They looked at the ostensive head-touch demonstration for 99.67% of its duration ($M = 18.26$ s, $SD = .19$ s), and at the interaction for 99.92% of its duration ($M = 11.99$ s, $SD = .02$ s). They looked at the first action event for 99.64% of its duration ($M = 11.04$ s, $SD = .16$ s), and at the second action event 100% of its duration. Half of the infants saw the lift-to-head action first, while the remaining half saw the elbow-action first. Which action was seen first did not affect infants' looking times at the two action events as the interaction of the order and the action type indicated no significant effect ($F(1, 14) = 1$, $p = .33$, partial $\eta^2 = .06$). Half of the infants saw agent A acting first while the remaining infants saw Agent B first. Which agent was seen acting first did not affect infants' looking times at the two different action events (F (1, 14) = 1, p = .33, partial η^2 = .07).

For test trials we again log-transformed the raw looking time durations. Data were collapsed across participant gender as preliminary analyses showed no significant interaction of this factor with test trials (F (1, 14) = .03, p = .87, partial η^2 = .002). There were also no significant interactions between which action was first seen and the test trials (F (1, 14) = 3.55, p = .08, partial η^2 = .2) and between which agent acted first and the test trials (F (1, 14) = .08, p = .78, partial η^2 = .005).

A repeated t-test on log-transformed looking times for the two actions revealed no significant difference, (t (15) = .18, p = .67, Cohen's d = .01). Figure 3.3a shows mean looking times at the two actions in seconds (for lift-to-head: M = 20.22 s, SD = 16.67 s; for elbow: M = 18.05 s, SD = 12.21 s). Figure 3.3b shows mean log transformed looking times for the two actions.

a. **b.**

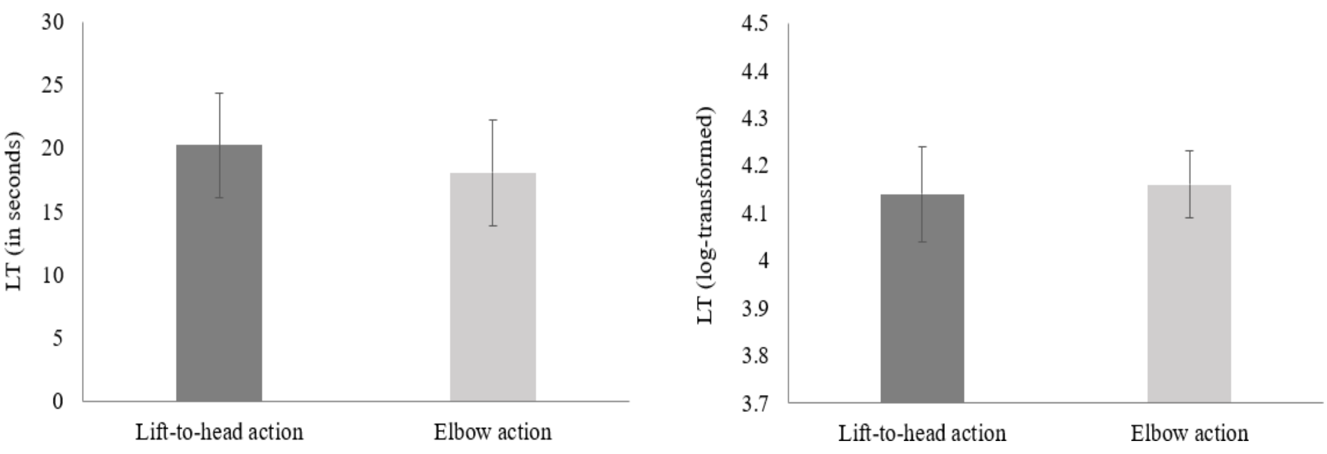

Figure **3.3a.** Mean looking times in seconds at the two actions for infants in Experiment 2. **3.3b.** Mean log-transformed looking times at the two actions. Error bars show ±1 SE.

3.4.4.1 Experiment 1 & 2

We pooled the data for Experiments 1 and 2 to explore whether the factor of age could account for the difference in looking time patterns of 14-month-old and 18-month-old infants at the lift-to-head and the elbow actions. For this purpose, we only included the infants who viewed the interaction in the congruent language condition ($N = 16$) in Experiment 1. As Table 3.1 shows 10 (out of 16) 14-month-old infants looked at the lift-to-head action longer whereas it was only five (out of 16) 18-month-olds who looked at the lift-to-head action longer. A Fisher's exact test indicated only a trend *in* the proportion of infants looking longer to the lift-to-head action between the two experiments ($p = .08$, one-tailed)

Table 3.1. The number of infants looking at the lift-to-head action longer in Experiments 1 and 2

	Looking longer to lift-to-head action	Looking longer to elbow-action	Total
Experiment 1	10	6	16
Experiment 2	5	11	16
Total	15	17	32

A repeated measure ANOVA with a between-subject factor of experiment also revealed no significant interaction between experiment and action type (F (1, 30) = 2.21, $p = .15$, partial $\eta^2 = .07$), and no significant main effect of action type (F(1, 30) = 1.54, $p = .22$, partial $\eta^2 = .05$) on looking times. Also, contrary to our prediction, the difference in infants' looking times at the two different actions did not correlate with age (r (30) = .003, $p = .98$).

3.4.5 Discussion

We hypothesized that unlike 14-month-olds, 18-month-old infants may look longer to the novel "elbow-action", in comparison to the "lift-to-head action" performed by the same-language speakers after viewing an ostensive head-touch demonstration by the informant. Given the possible reasons we hypothesized for why 14-month-olds may have looked longer at the more familiar test event in Experiment 1, we assumed that older infants could better encode the complex series of events presented to them, hence we expected that they may look longer at the novel test event as was originally hypothesized. This was not the case, however. Instead we found that older infants looked equally long at both test actions.

Following the argument concerning relative preference for novelty versus familiarity that we discussed in relation to the findings of Experiment 1 (given the complexity of the event sequence presented), we could either proceed to test even older infants using the same stimuli or investigate whether one of the test actions (the "lift-to-head") would elicit longer looking than the other test action (the "elbow-action") simply due to the different salience of the lift-to-head action over the elbow action. There could be several reasons why infants may have found the lift-to-head action inherently more interesting. First, the lift-to-head action involved more motion. Moreover, when the agent lifted the box to her forehead to bring about the effect, the agent's forehead and partially her face were lit up for two seconds due to the light emitted by the box. In contrast, for the agent who touched her elbow to the box the light-effect only illuminated the tip of her elbow (and not her face). This also could have made the lift-to-head action more salient in comparison to the elbow action.

Nevertheless, it is important to note that our looking time measurement started when each agent completed her action and lowered her hands to a resting position on the table at the two sides of the apparatus (see Figure 3.1e and Figure 3.1g). Hence there was no motion in the scene nor was there differential light effect on the agent's face during the test phase. Yet it can be still the case that such salient characteristics of the lift-to-head action had carried over till the end of the test phase, driving infants' attention selectively. If 14-month-olds' longer looking at the lift-to-head action is driven by a low-level process (such as the relative salience of the lift-to-head action over the elbow-action) in Experiment 1, then a baseline experiment with no initial ostensive head-touch demonstration and without differential language information could be expected to reveal the same pattern of findings (i.e. longer look at the lift-to-head action). If an older age group also showed a baseline tendency of looking longer at the lift-to-head action, then this would suggest that in fact in Experiment 2 infants increased their looking at the elbow-action, in line with our initial prediction for Experiment 2. On the other hand, if older infants look equally at the lift-to-head and elbow actions in a baseline experiment without any prior ostensive demonstration of the original head-touch means action, then it is highly likely that they showed no sensitivity to the differential degree of similarity of the lift-to-head action to the original head-touch action. Our baseline experiments aim to test for these possibilities.

3.5 Experiment 3: Baseline conditions

Experiment 3 was designed a posteriori as a baseline study to explore whether infants in the two age groups tested found the lift-to-head action more salient in comparison to the elbow-action. For example, if 14-month-olds show longer looking to lift-to-head actions in a condition where they see no ostensive head-touch demonstration, then the longer looking times by the 14-month-olds to this action in Experiment 1 would simply reflect their selective attention to the action event they found more interesting. If 14-month-old infants look at the two actions equally long, then it is still a possibility that in Experiment 1 they took into account the ostensive head-touch demonstration they viewed initially and this sub-goal choice by the informant guided their expectations about how other agents belonging to the same social group as the informant should act on the box. Similarly, if the relative salience of lift-to-head action was a factor influencing the older infants' looking time responses as well, then the null findings we reported in Experiment 2 would in fact suggest a weak modulatory effect of ostensive head-action demonstration on 18-month-olds' interpretation of the two different test actions. Based on these considerations, we decided to run two baseline studies for each age group.

3.5.1 Participants

For the younger age group, participants were 16 infants (9 females). Mean age of the participants was 441.69 days (14 months 15 days) (range: 428-453 days; 14 to 14;27 months, *SD* = 7.62 days). An additional 13 infants were tested but excluded from the final sample due to experimenter error (*N* = 2), or inattentiveness/fussiness (*N* = 6). We further excluded four infants because they were looking at the test events less than 2 seconds in total and additionally one infant because he was looking at each testing event for 60 seconds (i.e. maximum time allocated for the test events).

For the older age group, participants were 16 infants (7 females). Mean age of the participants was 588.25 days (19 months 9 days) (range: 554-602 days, 18 to 19;23, *SD* = 13.77). An additional 4 infants were tested but excluded from the final sample due to experimenter error (*N* = 1), or fussiness (*N* = 2). We further excluded one more infant because he was looking at one test event less than 2 seconds in total.

3.5.2 Procedure and stimuli

The stimuli were the same as in the Experiment 1 except that infants in this baseline group did not view the initial ostensive head-touch action demonstration. Rather the first scene infants viewed started when two agents entered the scene and interacted with each other by exchanging an object. We dubbed this interaction video in FinalCut Express 4.0 software by muting the speech and adding infant-friendly melody to keep infants' attention on the screen. Doing so allowed us to remove the language information from the interaction, which had been a crucial manipulation, both in Experiments 1 and 2. This non-verbal interaction video was 12 seconds long and using the PRAAT software we ensured that the melody accompanying the video was matched to the decibel intensity of the voices of the two agents in previous experiments. This non-verbal interaction scene was followed by the two action events as in the case of Experiment 1. An attention-grabbing animation (black and white checkerboard moving in the center of the screen while making a ringing sound) was played before each movie. The procedure was otherwise the same as in Experiment 1. Which actor displayed which action, and which action was carried out first was counterbalanced across participants.

3.5.3 Coding

We coded infants' looking behaviors toward the screen frame-by-frame offline. Our coding criteria were the same as in the previous experiments. For the youngest age group, 31% of the looking time data (*n* =5) were coded by a coder naïve to the predictions of the study. Inter-coder reliability was high for action test trials (r = .99 for two test actions collapsed, mean absolute difference between the codings for the test actions was 88 ms, *SD* = 143.36 ms). For the oldest age group 31% of the looking time data (*n* =5) were also coded by a coder naïve to the predictions of the study. Inter-coder reliability was high for the action test trials (r = .99 for two test actions collapsed, mean absolute difference between the codings for the test actions was: 92 ms, *SD* = 88.54 ms).

3.5.4 Results

3.5.4.1 Fourteen-month-olds

All infants viewed at least one action display for each manner variant actions. They looked at the first action 97.2% of its duration (*M* = 10.77 s, *SD* = .7 s), and at the second action 100% of its duration. Half of the infants saw the lift-to-head action first, while the other half saw the elbow action first. Which action was seen first, did not affect infants' looking times at the two action events as the interaction of the order and the action type indicated no significant effect (F (1, 14) = .81, p = .38, partial η^2 = .05). Half of the infants saw agent A acting first while the remaining infants saw Agent B first. Which agent was seen first did not have an effect on infants' looking times at the action events (F (1, 14) = .81, p = .38, partial η^2 =.05).

For test trials we log-transformed the raw looking time durations. The data were collapsed across participant gender as preliminary analyses showed no significant interaction of this factor with test trials (F (1, 14) = .18, p = .67, partial η^2 = .01). There were no significant interactions between which action was first seen and the test trials (F (1, 14) = .15, p = .7, partial η^2 = .01) or between which agent acted first and the test trials (F (1, 14) = .06, p = .81, partial η^2 =.004)

A repeated t-test on log-transformed looking times for the two actions indicated a statistically significant difference, (t (15) = 2.39, p = .03, Cohen's d = .63). Figure 3.4a shows the mean looking times at the two actions in seconds (for lift-to-head: M = 17.54 s, SD = 12.42 s; for elbow: M = 12.17 s, SD = 7.4 s). Figure 3.4b shows mean log-transformed looking times for the two actions.

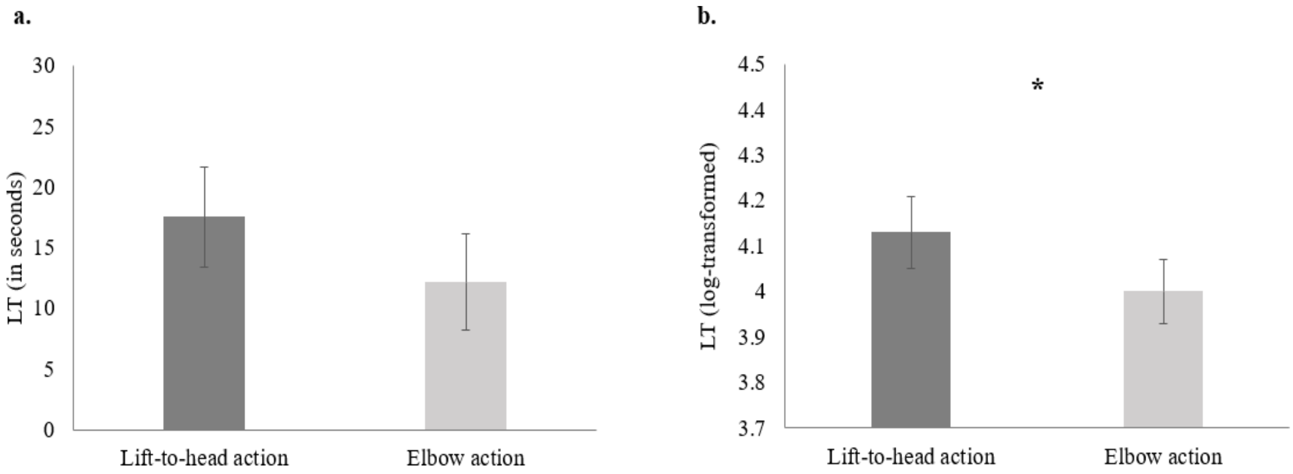

Figure 3.4a. Mean looking times in seconds at the two actions for 14-month-old infants in the baseline condition. **3.4b.** Mean log-transformed looking times at the two actions for 14-month-old infants in the baseline condition. Error bars show ±1 SE. Asterisk marks a statistically significant difference.

3.5.4.2 Older age group

All infants viewed at least one action display for each manner variant actions. They looked at the first action 97.2% of its duration (*M* = 10.77 s, *SD* = .87 s), and at the second action 100% of its duration. Half of the infants saw the lift-to-head action first, while the other half saw the elbow action first. Which action was seen first, did not affect infants' looking times at the two action events as the interaction of the order and the action type indicated no significant effect (F (1, 14) = 2.09, p = .17, partial η^2 = .13). Eight infants saw agent A acting first while the remaining infants saw Agent B first. Which agent was seen acting first did not affect infants' looking times at the two action events (F (1, 14) = .16, p = .69, partial η^2 = .01). For test trials data were collapsed across participant gender as preliminary analyses showed no significant interaction of this factor with test trials (F (1, 14) = .81, p = .38, partial η^2 = .05). There was no significant interaction between which agent acted first and the test trials (F (1, 14) = 3.67, p = .08, partial η^2 = .2). There was also no significant interaction between which action was seen first and the test trials (F (1, 14) = 3.06, p = .1, partial η^2 = .18).

A repeated t-test on log-transformed looking times of infants for two actions indicated no statistically significant difference, (t (15) = -.31, p = .97, Cohen's d = .02). Figure 3.5a shows mean looking times at the two actions in seconds (for lift-to-head: *M* = 18.18 s, *SD* = 12.4 s; for elbow: *M* = 16.47 s, *SD* = 9.6 s). Figure 3.5b shows mean log transformed looking times for the two actions.

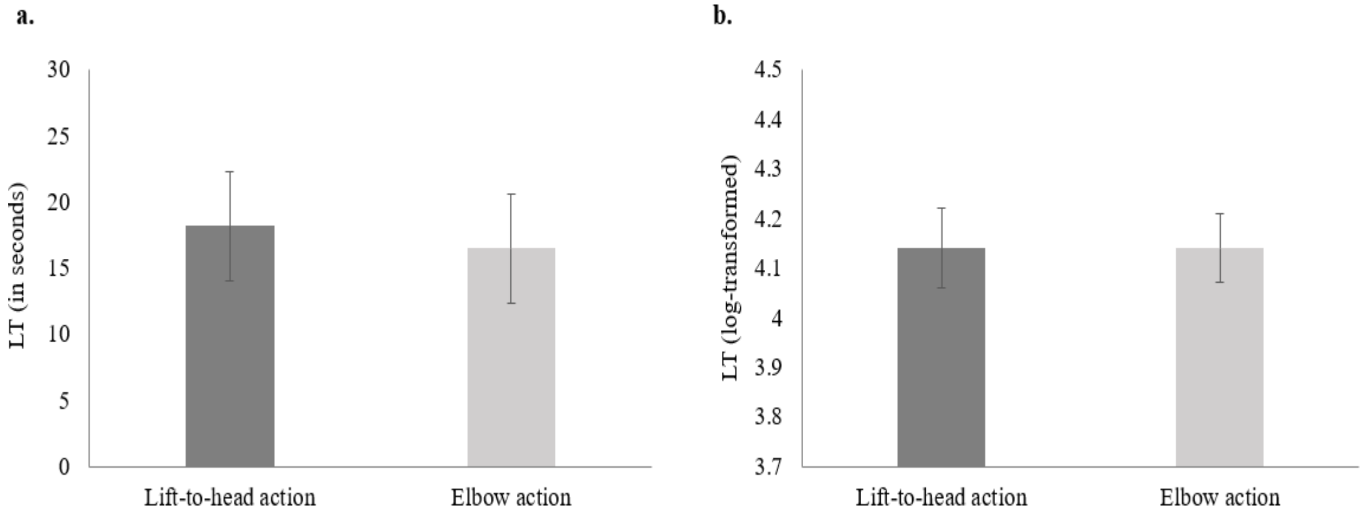

Figure 3.5a. Mean looking times in seconds at the two actions for older infants in the baseline condition. **3.5b.** Mean log-transformed looking times at the two actions for older infants in the baseline condition. Error bars show ±1 *SE*.

3.5.4.3 Experiment 1 and younger age-baseline group

A repeated measures ANOVA with a between-subject factor of condition (with three levels of congruent, incongruent, and baseline condition with younger age group) revealed a statistically significant main effect for action type on 14-month-old infants' looking time durations (F (1, 45) = 7.17, p = .01, partial η^2 = .14). However, there was no interaction between the condition (with three levels) and the action type (F (2, 45) = .21, p = .81, partial η^2 = .009). The mean looking time for the lift-to-head action was 19.03 s (*SD* = 14.15 s) while the mean looking time for the elbow action was 14.25 s (*SD* = 12.25 s).

3.5.4.4 Experiment 2 and older age-baseline group

A repeated measures ANOVA with a between subject factor of condition (with two levels of congruent condition and the baseline condition with older age group) revealed no statistically significant main effect for action type on older infants' looking time durations (F

(1, 30) = .01, p = .92, partial η^2 = .00). There was no interaction between the condition and the action type (F (1, 30) = .02, p = .89, partial η^2 = .001). Infants in the two conditions looked equally long to both sub-efficient means actions (M_2 = 20.23 s, SD =16.68 s for lift-to-head action; M_2 = 18.06 s, SD = 12.21 s for elbow action in experiment 2; M_{baseline} = 18.18 s, SD = 12.4 s for lift-to-head action; M_{baseline} = 16.48 s, SD =9.6 s for elbow action in the baseline).

3.5.5 Discussion

Fourteen-month-old infants looked longer to the lift-to-head action in contrast to the elbow action in a posteriori designed baseline study that included no ostensive demonstration of the leaning-on-the-box action. This finding reveals that the relative salience of one type of action over the other could have been a factor driving 14-month-old infants' looking behavior at the two test trials. This finding further documented that prior ostensive demonstration of the head-touch action had no influence on infants' looking time behaviors at the different test events. Rather it seems to be the case that 14-month-olds' longer looks to the lift-to-head action than to the elbow action stemmed from the differential salience of mean actions tested in this paradigm. Since the lift-to-head action involved more motion, and it also caused a light effect on the forehead of the agent performing it, 14-month-old infants might have attended to it more than at the elbow action. This was the case even though we started to measure infants' looking time after the agents ended their respective sub-efficient mean displays, with their hands resting at the sides of the lightbox. It is still a possibility that the effect of differential salience of the two actions observed during the action events carried over until the end of the test phase, leading 14-month-old infants to still pay more attention to the lift-to-head action in comparison to the elbow-action. On the other hand, exactly as the 18-month-olds in Experiment 2, a new group of older infants in the baseline condition looked equally at the lift-to-head action and the elbow-action. The differential salience of the actions that resulted in 14-month-old's longer baseline looking at one type of action than at the other did not have a similar influence on older infants' looking times.

3.6 General discussion

In Experiments 1 and 2 we showed that neither 14-month-olds nor 18-month-olds expected the particular sub-efficient means action (i.e. head-touch action on touch sensitive light-box) that was ostensively demonstrated to them by a Hungarian informant to be shared among the members of the same social group. Specifically, our hypothesis was that when infants view an ostensive demonstration of a sub-efficient manner (e.g. leaning over the box to touch the light-box with the forehead) they will expect this sub-efficient manner to generalize over other social group members, and will expect agents sharing the same language with the ostensive informant to behave similarly (i.e. by performing a head-contact action). However, 14-month-olds looked longer when they had seen another agent performing a "lift-to-head action" in comparison to when they had seen another agent doing an "elbow action". In other words, we could not elicit a violation of expectation in infants upon seeing the "elbow action". Eighteen-month-olds looked equally at the lift-to-head and the elbow actions. Our baseline studies mirrored the pattern of findings we had in the congruent language condition for both age groups: 14-month-old infants looked longer at the lift-to-head test event in comparison to the elbow test event whereas 18-month-olds looked at the two test events equally. It was relative salience of the lift-to-head action over the elbow action that drove 14-month-olds', but not 18-month-olds' attention to the different action displays. So, the ostensive demonstration had no influence on infants' looking time behaviors.

Active behavioral paradigms documented that 14-month-olds do not only emulate the goal but also the sub-goals demonstrated to them in ostensive contexts. For example, after viewing an ostensive demonstration of the informant leaning on the light box to bring her forehead into contact with its surface infants were likely to press their chins, cheeks, lips, and other parts of their head against the surface of the box in order to activate it (Kiraly et al., 2013). Similarly, when the light box apparatus was small enough to be easily lifted by hand infants brought the box to their foreheads using their hands in order to activate it, instead of leaning over to touch the box with their foreheads as it was shown to them (Chen, Kiraly, & Gergely, 2012). Based on infants' responses in behavioral paradigms showing their tendency to emulate the demonstrated sub-goal, we assumed that they can perceive the two actions of lean-on-box-with-head and lift-the-box-to-head as guided by the same sub-goal as both alternative actions resulted in *contact between the box and the forehead*. Since we measured the looking time behavior of infants by having the last still frames of the action events on the screen, we did not take into account that perceived relative salience of one action over the other would carry over to the test phase creating a confound that could influence the looking times of infants, especially when they were younger.

Our findings for both Experiment 1 and 2 suggest that the link between the ostensively demonstrated target action and its emulative version in a third-party observation context was not detected by the infants in either age groups (i.e. when they view one agent "emulating" another agent). There could be several reasons for this. First of all, unlike Liberman and her colleagues (2018) we did not provide infants with an active exploration phase where they could have interacted with the lightbox before the looking time study. Infants might need to realize the end goal of the action (i.e. light effect) by the most efficient means, i.e., with their hands, to be able to encode the demonstrated head-action in the teaching event as the sub-goal that is the relevant and novel piece of information in the event sequence. After all, infants did not see this apparatus before, and not only the sub-goal performed but also the target apparatus itself and the goal achieved were all novel for them. Indeed, in all behavioral studies using the head touch paradigm, infants tended to activate the lightbox with their hands first (Gellen &

Buttelmann, 2017; Gergely et al., 2002; Kiraly et al., 2013). A self-exploration phase might have allowed them to represent the overall goal of the action as lighting up the box and to infer that the demonstrator could have illuminated the box with her hand if she was only demonstrating the function of the box. This self-exploration phase might have been a factor boosting infants' understanding that there is a special reason behind the demonstrator's sub-efficient means act. This in return could have helped them to encode the sub-efficient means demonstration as a relevant sub-goal to acquire.

Secondly, while two agents acted on the light box with their respective means the ostensive demonstrator attended both agent's actions with a friendly expression. This alone might have sanctioned both action-means as equally acceptable. Additionally, the two agents on the sides were viewed as engaging in a prosocial interaction with one another (i.e. exchanging a a paper tissue). The mere fact that these two agents were prosocially engaged with each other first but then later displayed two distinctively different means might have been confusing for infants (see Liberman et al., 2018), contributing to the complexity of the sequence of events infants were presented with.

Third, presenting infants with a brief ostensive context in which an ostensive informant addresses them and shows a sub-efficient means of activating the box might not be salient enough of a cue to help infants to encode the action as a socially relevant sub-goal especially given that both the apparatus and the means adopted were novel. Previous research has found that when naïve observers view multiple agents acting in the same way, and sometimes in synchrony with each other, then these cues help them to interpret the action from a ritual stance, enabling them to faithfully imitate the causally opaque components of the action sequence (Hermann et al., 2013; DiYanni, Corriveau, Kurkul, Nasrini, & Nini, 2015). Clearly not even the 18-month-olds in our study encoded the initially demonstrated ostensive sub-efficient action manner as a socially stipulated sub-goal, as they did not expect other parties to act on the same apparatus with an equivalent emulative action variant. Future research should use multiple agents as demonstrators, and perhaps should also exploit the cues of synchrony, in addition to ostensive signals, while aiming to induce an interpretation of ritual stance for sub-efficient means actions. For example, there could be two agents on the scene speaking the same language as the infants. After familiarizing infants to the identical sub-efficient means action these two agents perform in activating the light box in synchrony (e.g. Hungarian agents would always lean on the box to touch their forehead to its surface), infants then could be shown two test events: one that is consistent with their expectations (e.g. another Hungarian agent performing the same sub-efficient means as the agents with whom infants were familiarized), and another that would be inconsistent with their expectation (e.g. another Hungarian agent performing an elbow touch action). While we purposefully avoided using the exact same means action display (e.g. lean-on-the-box with the head) during the test phase as the one that had been ostensively demonstrated to infants (anticipating that infants would look longer at the elbow action just because of its perceptual novelty) future research might use the same sub-efficient mean action in the test as it was demonstrated to infants in an ostensive demonstration context . If our hypothesis is right, infants tested in a similar design as proposed above, but this time with foreign-language speakers during the test, would be surprised upon viewing the foreign actor’s use of the sub-efficient means that had been demonstrated earlier by the Hungarian speaking informants.

3.6.1 Final remarks

Given that looking time findings might not always lead to convergent results than what are found when using active paradigms (e.g. rational imitation looking time studies, Elsner, Pfeifer, Parker, & Hauf, 2013; Buttelmann, Schieler, Wetzel, & Widmann, 2017) we are hesitant to completely ignore the theoretical significance of our hypothesis. It remains an interesting question for future research whether or not infants can infer third-party relations, and recognize group members, not only by observing others faithfully imitating each other (as Powell & Spelke, 2018a already showed) but also observing others *emulating* each other. Yet given that the cue of ritualistic action in recognizing the same group members is characterized by inflexibility and stereotypy of the particular action (Boyer & Lienard, 2006), it might be the case that if a third party does not faithfully imitate the ritualistic action performed by the same-group member, then even if she makes use of the same unusual body part in performing the ritual act the emulated-new action is not interpreted anymore as an equivalent ritualistic act. After all, even if two different versions of the sub-efficient means actions are guided by the same sub-goal (e.g. head-contact), arbitrariness of such actions is what makes them good candidates for conventional practices in cultural communities. Given that it is highly unlikely for a foreigner to discover by self-exploration exactly the same arbitrary mean action that is shared knowledge within a different cultural community that she or he is not a part of, naïve learners might expect two individuals from the same community to perform sub-efficient mean actions exactly in the same way. Hence any deviation from this faithful performance would indicate that the individuals do not have the same cultural common ground. To test this, following a similar design as I proposed above, infants could be presented with two agents from the same linguistic group during the test: one faithfully adopting the way her group members previously performed (lean-on-the box with the head), and another emulating this sub-goal (bring-the-box to head). I would expect infants to be surprised when the agent emulated the sub-goal her group members previously performed. In the case of agents belonging to a different linguistic group, I would expect infants to be surprised when the agent faithfully imitated the different group members' manner of action. It might be the case that infants expect third parties that are from the same social group to only faithfully imitate each other's manners; and emulative behavior is not a cue for recognizing the same group members.

Chapter IV. Preverbal infants’ sensitivity to inefficient and efficient means actions shared among the social group members

4.1 Introduction

Infants expect agents to act efficiently (Gergely et al., 1995; Csibra, 2003), but they also expect agents belonging to the same group to behave alike (Powell & Spelke, 2013). However, how these two early emerging sensitivities interact with and related to each other in governing infants’ expectations about others’ actions has not received much empirical scrutiny. Infants might show sensitivity and flexibility in representing agents whose actions are governed by instrumental principles of rationality, on the one hand, and by the shared conventions of the social groups they belong to, on the other. Especially when infants observe agents who belong to the same social group to pursue their goals through performing sub-efficient mean actions, they might show readiness to look for and rationalize the sub-efficient manner of means performed by reference to external reasons (such as belonging to the same social group) to account for why those agents chose to bring about their goal that way. Relying on cues about the social group that the agent belongs to might provide relevant information for infants to interpret the reason behind the agents’ performance of the sub-efficient mean actions to achieve their goal. This would enable infants to represent that the agent possesses and employs the shared action repertoires of their social group, which would constitute the foundations for understanding group-specific norms, traditions, conventions, and culture-specific rituals that would account for their choice of manner of means actions over-and-above the principle of efficiency driving goal-directed instrumental acts. However, this would also require infants to be able to suspend their expectation of efficiency of means actions under certain social contexts. This chapter aims to investigate how these two kind of expectations for efficient goal-directed acts on the one hand, and for shared group-specific behaviors on the other, interact, if they do at all, in guiding infants' inferences about expectable actions of agents who belong to the same social group.

4.1.1 Representing goal-directed actions of agents in social groups

According to the core knowledge hypothesis infants represent and guide their reasoning about various domains by relying on specific evolved systems of principles that form innate adaptations and are shared with other non-human animal species (Carey 2009; Spelke, 2016; Spelke, Bernier, & Skerry, 2013). One such domain-specific core knowledge system is dedicated to representing intentional agents. Infants identify agents through their self-propelled actions as entities who can cause their own behavior, who can contingently react to other entities' behaviors at a distance and can act intentionally on objects to bring about some change of state in their physical environment. Crucially, such goal-directed actions are bounded by the constraints of the physical environment within which infants represent the most efficient manner available to the agent in bringing about the intended goal state (Carey, 2009; Gergely & Csibra, 2003).

Apart from such an evolved system to represent goal-directed instrumental agency, Spelke (2016) has recently proposed another core system that is dedicated to representing the domain of "social beings". According to this hypothesis, representing entities as social beings who engage with one another on the one hand, and as instrumental goal-directed agents on the other, constitute two independent and separate adaptations of core knowledge systems. The primary property of instrumental agents is to purposefully act on objects to bring about an intended physical change in the environment, while the primary property of agents categorized as "social beings" is to perform actions to engage with one another in order to "express their attention, emotion and commitment" to each other (Spelke, 2016, p. 292). Identifying agents as "social beings" is based on recognizing their social behaviors which are often causally opaque, arbitrary and conventional behaviors that are directed towards their social partners without being performed instrumentally to bring about a physical change of state in the environment (Powell & Spelke, 2018a). This hypothesis suggests that infants can clearly differentiate intentional actions that signal and serve social affiliation with other "social beings" from goal-directed instrumental actions directed to physical objects with the instrumental aim to bring about a physical change in the object environment. Furthermore, this proposal claims that preverbal infants cannot conceive an intentional action of another agent as being both a social and an instrumental goal-directed action at the same time and that representing other entities as instrumental agents and as social beings might invoke different representational content in infants (Spelke, 2016).

However, this strong hypothesis appears to us questionable on the basis of evidence provided by previous studies which suggest that infants can represent other entities as both being instrumental agents and social beings, who act on objects in a goal-directed manner while also interacting with each other as social partners. For example, empirical findings on infants' understanding of cooperation scenarios suggest that they represent other entities both as instrumental agents and as social beings at the same time who collaborate with one another to achieve a common goal and they can do so at least as early as 13-months of age (Henderson & Woodward 2011; Henderson et al. 2013; Vorobyova, Teglas & Gergely, 2017). In fact, a recent study by using a NIRS brain imaging methodology indicates that understanding of coordinated actions in the pursuit of joint goals can be demonstrated as early as 9-month of age (Begus, Curioni, Gergely, & Knoblich, 2018). This finding implies that young infants already before their first birthday can understand and attribute a goal not only to individual instrumental actions but can also interpret the coordinated actions performed simultaneously by two social agents and attribute a shared joint goal to them. At 16 months infants also make inferences about the affiliative status between two characters based on the kind of sub-efficient mean actions they were observed to perform in pursuit of bringing about a physical change in an object manipulated (Liberman et al., 2018). Thus, the available empirical evidence indicates that infants might be able to represent self-propelled entities both as instrumental agents and as social beings (as “social agents”) and can integrate social and instrumental actions performed by them from a relatively early age on. Therefore, based on the kind of evidence reviewed above, we assume that young infants’ inferences about other agents’ actions could be guided both by instrumental and social factors at the same time. Specifically, we hypothesize that infants can interpret the observed actions of social agents as being guided and constrained by two different basic assumptions: the principle of instrumental rationality and the shared movement repertoire that characterize the social group the agents belong to and affiliate with.

4.1.2 Do infants expect social agents to act alike?

The burgeoning literature on infants’ action interpretation documents that infants expect agents to act efficiently to achieve goals (Csibra, Gergely, Bíró, Koós, & Brockband, 1999; Gergely, Nadasdy, Csibra, & Biro, 1995; Biro, Verschoor, & Coenen, 2011, Csibra, 2003; 2008a). But do infants also expect agents to act alike as other social agents do? A recent study investigating infants’ inferences about the acts of social characters has relied on infants’ social group representations to document that infants use group membership information to guide their inferences about how agents would behave (Powell & Spelke, 2013). In this series of experiments infants were presented with two different social groups both of which consisted of animated characters possessing facial features and a capacity for self-propelled motion. The different groups were characterized by the cues of identical appearance (orange stars versus purple trapezoids), spatial proximity, synchronous and identical dance-like movements and the identical sounds produced by their members. After having been familiarized to two different groups each consisting of three characters, infants watched two characters from the same group jumping up and down on a platform and the other two characters from the other group performing a different action of sliding from right to left on the same platform. Twelve-month-old infants looked longer when the third member from each group behaved inconsistently with the other members of their groups (i.e. jumping when the other group members slid or sliding when the other group members jumped). In other words, infants expected group members to move alike (Powell & Spelke, 2013).

Infants' propensity to represent the characters presented to them in social groups seems to guide their expectation about how novel agents belonging to those groups should behave like. Importantly, however, this tendency was demonstrated by Powell and Spelke (2013) by employing intransitive actions. The agents simply jumped or slid on a platform and these movements were not directed towards an external object in the environment. Infants watching such intransitive actions expected the same-group members to move in the same way as the other members of their group. Follow-up experiments by Powell and Spelke (2013) also documented that infants as early as 7-months expected the same-group members to act on the same goal object (i.e., to land on the same box) and were surprised when the group member acted on a different goal object than the one its group members had acted on before (i.e., when it landed on a different box). Critically in these studies members from both groups circled around the target object before landing on it, hence they performed the same mean actions to achieve their goals as did the other members of their group before.

This leaves unanswered, however, whether infants would expect two different social groups to perform different means in pursuit of the same goal. After all, in human cultural practices agents use different tools, different symbols and different means actions, even if these are performed with the intention to serve similar goals (examples range from using chopsticks versus fork and knife while eating, from putting a hat on or taking it off while going into a place of worship). Most of the distinctive ways through which different cultural groups achieve their similar goals are distinguishable from each other due to the particular characteristics of the specific manner of actions that the different cultural groups endorse as their group-specific means. For example, culturally relevant means adopted by a social group are often not the most efficient (optimal) way of bringing about the goal (you can more efficiently eat food by using your hands instead of using a chopstick or a fork), and sometimes the culture-specific mean actions are not even causally necessary to achieve the goal (neither putting your hat on nor putting it off is a causally enabling condition to allow you to physically enter either a synagogue or a church). Such sub-efficient and arbitrary mean actions, however, might be highly informative for the observer to identify and represent these actions as being governed by culture-specific conventional practices that are characteristic of the particular social group the agents belong to. Furthermore, the characteristic manner in which these actions are performed provide salient and relevant information to signal and identify membership in a specific social cultural group and so they could be usefully encoded and represented as relevant information that specify the *normative and conventional manner* of how one *"ought to" act like* if one is a member of a particular social community (Gergely, 2013; Kiraly 2009b; Kiraly et al., 2013).

For this reason, infants should be able to put their expectation of instrumental efficiency of goal attainment on hold when interpreting the inefficient way of attaining a goal in certain social contexts. For example, being presented with multiple agents from the same social group acting inefficiently yet purposefully in achieving the same end state might help infants to represent the inefficient means action as an essential and relevant element of the action sequence, that is to be shared among the same-group members. On the other hand, any agent can pursue the given goal efficiently independent of the social group she or he belongs to, and - given the dominance of the efficiency principle in realizing instrumental goal-directed actions - one could assume that young infants would expect other agents to pursue their goals efficiently. If this were so, then infants should expect other agents to approach their goals efficiently given the physical constraints of their environment, and that they should ignore the inefficient manner of action particular social group members performed to achieve their goal. However, if infants showed special sensitivity to expect social group members to align their behaviors with each other, then they should be able to relax their efficiency criterion and should expect that members of the same social group to act alike even when their actions appear to be inefficient or unnecessary to achieve their goal. In this chapter our aim is to test this hypothesis in a 2 x 3 design that pits social group membership against the efficiency of the means actions performed as can be seen in Table 4.1. Note that since we did not have a theoretically strong hypothesis for all the six experimental conditions that this study design yielded, we did not carry out all of them.

	Individual	Ingroup	Outgroup
Familiarized to efficient actions	Experiments 1a & 1b	Experiment 3	Experiment 4
Familiarized to inefficient actions		Experiment 2	

Table 4.1. 2 x 3 experimental design outlining the five experiments carried out in Chapter 4

All five experiments we report below used a violation-of-expectation looking time paradigm, and all were pre-registered on Open Science Framework before the data were collected. Infants first observed a group induction event similarly to the procedure used in the study by Powell and Spelke (2013). These group induction events presented three agents belonging to the same group (labelled as “ingroup” agents), and a fourth agent belonging to different group (labelled as the “outgroup” agent). Group membership was indicated by the very same cues used in the original procedure of Powell and Spelke (2013). Ingroup agents were orange triangles while the outgroup agent was a blue circle (counterbalanced); ingroup agents moved synchronously with their own group members while the outgroup agent moved alone; ingroup agents produced a sound identical of their group while the outgroup agent produced a different sound; ingroup agents always appeared next to and in close proximity to each other, while the outgroup agent was positioned further away from them. Note that only in Experiment 1b were infants not presented with this group induction event.

In all five experiments infants then viewed action familiarization events depicting four agents stacked up one above the other on the left side of the scene (showing the identically shaped- and colored- ingroup agents in close proximity to each other while the differently shaped- and colored- outgroup agent was further away from them). Note that in Experiment 1b, there was only one agent in the scene. During the action familiarization phase infants saw either an individual ingroup member (Experiment 1a) or an individual agent (Experiment 1b) or two ingroup members (Experiment 2 – 4) acting towards the goal object that was on the right side of the scene. Critically, during familiarization, there was a long barrier in the scene separating the agent(s) from the goal object (Experiments 1, 3, 4). The agent(s) moved over the barrier in order to obtain the goal object that was resting on the other side. The height of the barrier justified the moving-above-the-barrier action of the agent(s), rendering it an efficient goal approach. In Experiment 2, the barrier was short allowing the agents to take a straight path under it to reach the goal. In spite of this, the agents moved along a path above the barrier to reach the goal object. Thus, the length of the barrier did not justify the moving-above-the-barrier action of the agents, making the actions inefficient in Experiment 2.

During the test phase, in all five experiments, the barrier was always short and did not reach all the way down. Infants viewed either the same agent who acted during familiarization (Experiments 1a & 1b) or a novel ingroup (Experiments 2, 3) or an outgroup agent (Experiment 4) approaching the goal object in an efficient manner (by taking the straight path towards the goal) or approaching it through an inefficient manner (by moving-above-the-barrier) in alternate test trials. We measured infants' looking time to the test trials in all of the five experiments. Additionally, we also registered infants' pupil dilation as a further dependent measure in an exploratory analysis.

Experiment 1a aimed to verify that our stimuli induce teleological inferences about the efficiency of goal approach in 11-month-old infants. The group induction event was not relevant to this experiment as infants viewed the same agent acting during the familiarization and the test events. However, in Experiment 1a we also presented the same sequence of events as we used in Experiments 2-4 in order to make sure that the length of the events infants saw before the test was equal across the experiments.

4.2 Experiment 1a. Individual efficiency

The aim of Experiment 1a was to verify whether the particular types of unfamiliar and abstract entities that we designed to use as our stimuli across our planned series of studies, could successfully induce in 11-month-old infants the attribution of instrumental goal-directed agency based on observing such a single entity perform a self-propelled behavior directed to contact a distal goal object on the other side of the screen. In particular, we wanted to demonstrate that observing such an abstract entity's behavioral act to approach and reach a distal object would be sufficient to induce in 11-month-olds teleological inferences about the expected efficiency of the agent's goal-directed actions (as demonstrated by numerous previous studies, e.g., Gergely et al., 1995; Csibra et al., 2003). To make the results of Experiment 1a comparable to the rest of our planned experiments, we first presented group induction event to 11-month-olds.

During the *group induction* the infants observed four unfamiliar entities perform various behaviors while each emitted a specific sound. Previous research by Powell & Spelke (2013) documented that infants from the age of seven months are sensitive to and can rely on a number of behavioral and featural cues to represent several interacting entities as agents who belong to the same social group. The cues Powell & Spelke identified as inducing in infants the representation of the entities as agents belonging to the same social group involved the identical featural appearance of the entities (including having eyes), their spatial proximity, their synchronously performed and similar intransitive (dance-like) movements, and their contingent emission of an identical sound. Therefore, in Experiment 1a we used the very same cues as Powell & Spelke (2013) described: during the initial *group induction phase* infants saw three of the four entities with eyes exhibit all of the above featural, spatial, and behavioral cues (labelled here as “ingroup” members), while the fourth entity behaved and looked differently than the other three and produced a different sound than what the other three produced. According to Powell & Spelke’s hypothesis, this fourth entity would not have become categorized as an agent belonging to the same social group as did the other three (so we label this agent here as an “outgroup” agent).

After having presented group induction scenes to infants using the above cues, we further familiarized them with efficient goal-directed actions performed by one of the ingroup agent who approached the goal object by moving above and over the long barrier to obtain the goal object that was placed at the other side of the barrier. During the test trials the barrier was shorter (its lower part did not reach all the way down) and so it did not block direct access to the goal object. On alternate test trials infants watched as the same ingroup agent (who had retrieved the goal object in an efficient manner during familiarization) either approaching the goal object in a different, but nevertheless efficient pathway by taking the most direct (straight-line) route to reach the goal (efficient approach trial) or following the same path of going over the barrier as during familiarization even though during test this approach route was not the most efficient alternative available (inefficient approach trial). Note that the information about the agent’s group membership presented during the group induction was not relevant in this experiment, yet it was still presented to infants to make sure that the length of both the group induction and action familiarization phase was identical in all experiments and so differences in the amount of looking induced by the two types of test trials could not be the result of differences in the relative length of the familiarization stimuli.

We predicted that infants would look longer at the inefficient test trials than the efficient test trials, replicating previous findings which had demonstrated infants' expectations of efficiency of goal-approach (Gergely et al., 1995; Csibra et al., 2003). We also registered infants' pupillary response as pupil dilation has been shown by previous studies to indicate increased cognitive load (Laeng, Sirois, & Gredeback, 2012; Hepach & Westermann, 2016). We therefore predicted that inefficient test trials would evoke greater pupil dilation than efficient test trials due to the differential processing demands associated with expected (efficient) versus unexpected (inefficient) test trials. Given the close link between increased attention and pupil size (Laeng et al., 2012) we expected greater pupil size to reflect infants' increased attention to inefficient actions in comparison to efficient actions during the test trials.

4.2.1 Participants

Participants were 24 infants (12 females). Mean age of the participants was 332.54 days (10 months 27 days) (range 321-351 days; 10;16 to 11;16, *SD* = 8.66 days). An additional 25 infants were tested but excluded from the final sample due to mother interference (*n* = 2), inattentiveness (*n* = 1), fussiness (*n* = 4), looking-times of less than 2 seconds per trial (*n* = 5), and premature termination of a trial resulting from a lack of eye-tracking data for 2 consecutive seconds rather than looking away from the screen (*n* = 13). Please see Appendix 2 for the reliability codings for eye tracker registration errors.

Participants were recruited from a database of parents who volunteered to participate in developmental studies. Any willing parent was eligible to participate in the study with their infants as long as the infant had no sight or hearing problems, was not pre-term born, and she or he was 11 months old (+/-15 days) (i.e. between the age of 10 months 15 days, and 11 months 15 days). Each parent gave his or her written informed consent for the study, and the procedure was approved by the United Ethical Review Committee for Research in Psychology, Hungary, and conducted in accordance with the Declaration of Helsinki. Infants were given a toy or other small gift at the end of the study independent of their performance. Note that the same selection criteria were applied to all the participants who were included in the five experiments reported in this chapter.

4.2.2 Procedure

Infants were tested in a dimly illuminated and soundproof room. An Eyelink 1000 Plus remote eye-tracking system was used (SR research) to record infants' looking time behaviors and to collect their pupil dilation data. Infants sat on their parent's lap approximately 60 cm away from a 24-inch LED screen with the eye tracker camera attached underneath. The eye-tracker camera recorded the reflection of an infrared light source on the infant cornea from the right eye only, at the frequency of 500 Hz. We placed a small black and white sticker on infant's forehead to aid the eye-tracker in keeping the infant's gaze position on the screen. The parents were instructed to keep their eyes closed and to remain silent during the whole procedure. Two experimenters were sitting in the control room that was divided from the testing area with a thick black curtain where the infant and the parent were seated. The calibration procedure (with a nine-point calibration sequence) and the stimulus presentation were carried out using Experiment Builder software in an infant-controlled fashion (i.e. if the infant looked away from the screen for two consecutive seconds or more during the test event then the next test event was shown automatically by the software). Infants watched the events in the following order, yielding six test trials in total:

- *Group induction* (three dance events for each group of agents)

- *Action familiarization* (four action events)

- *Test* (two test trials)

- *Group reminder* (one dance event for each group of agents) & *second action familiarization* (two action events)

- *Test* (two test trials)

- *Second group reminder* (one dance event for each group of agents) & *third action familiarization* (two action events)

- *Test* (two test trials)

Infants were presented with an attention getter between each movie and between the test trials.

4.2.3 Stimuli

The stimuli consisted of short series of animation movies that were made using Adobe Animate CC software. In *the group induction phase* there were three characters with the same circular shape and the same color positioned close to each other on the scene (ingroup agents), and another character with a triangular shape and of a different color who was spatially further away from the other three agents on the scene (outgroup agent). We counterbalanced which geometrical shapes and color were the ingroup agents (half of the infants saw three blue circular agents as the ingroup agents and an orange triangle as the outgroup agent, while the other half saw three orange triangles as the ingroup agents and a blue circle as the outgroup agent). All these four entities were characterized by two agency cues: self-propelled motion and having eyes.

During *the group induction phase* either one of the ingroup agents or the outgroup agent was jumping up and down while producing a sound as an invitation for other group members to join the dancing event. Infants then watched the ingroup agents dancing in synchrony with each other as a group while making the same sound that was produced by the first agent as they were moving closer to and then away from each other in a circular motion (in the same manner as in the paradigm used by Powell & Spelke, 2013). This motion pattern was repeated three times resulting in three dancing events. The outgroup agent moved in the exact same pattern as the ingroup agents, but it performed the dancing movement alone, and produced a different sound than the one made by the ingroup agents (see Figure 4.1a). Ingroup agents and the outgroup agent took turns doing their dancing movements and which "group" started to dance first was counterbalanced across infants. There were six dancing events in total in the following counterbalancing order (with A referring to the ingroup agents' dance and B referring to the outgroup agent's dance: AB-BA-AB or BA-AB-BA). This *group induction phase* took 57.4 seconds in total.

The *group induction phase* was followed by *the action familiarization phase* in which infants first saw an object (a red star) sliding to the scene from the right side while emitting a sound (for the duration of .9 seconds). When the star stopped and stayed motionless on the right side of the scene, the ingroup agent, who was positioned at the bottom left, started approaching it. Since there was a tall barrier (reaching almost all the way down to the surface) between the agent and the goal object (the red star), the agent needed to change course when getting close to the barrier (to avoid collision) and continued moving upwards to go over the top of the barrier and then to descend on the other side to reach the goal object (see Figure 4.1b). After making contact with the object the agent turned and carried the goal-object back going over the barrier again in the opposite direction. This action event ended when the agent was at its starting position on the left side of the barrier, with the object on its left. We repeated each action event four times. Each goal-approaching action took seven seconds, while the full action including carrying the goal-object back to the starting position took 14.1 seconds. The total duration of *the action familiarization phase* was 62.3 secs.

During *the test phase*, the barrier that had separated the agent and the object during familiarization was shortened by 40% of its size from below (see Figure 4.1c and d). Its lower part was now missing thus forming no obstacle anymore between the approaching agent and the goal object on the other side. This rendered the agent's action of going over the top of the barrier unnecessary to reach the goal object (inefficient approach), as the agent could now approach the object more efficiently by following a direct pathway going under the barrier (efficient approach). *The test phase* consisted of two test events. One test event involved an efficient goal-approach, where the same agent who had acted during *the action familiarization phase*, approached the goal along a straight direct path. The other test event was inefficient, where the same agent who had acted during *the action familiarization phase* now approached the goal object by moving on the same path again going over the top of the barrier just as during the action familiarization phase.

The efficient and inefficient test events started the same way with a still frame of 1.5 seconds, later the goal-object slid to the scene from the right side and stayed motionless on the screen. At 2.4 seconds the agent started approaching the goal object along a direct straight path. At 7.1 seconds (henceforth "moment of divergence") the agent in the inefficient test event began to diverge from the direct motion trajectory and made a detour around going above the barrier while accelerating its motion. The agent was thus following the same pathway of going over the top of the barrier exactly as it did during the action familiarization phase. Both the efficient and the inefficient approach took the same time, and both actions ended as the agent moved the object back to the left of the scene to its starting position as it did during the action familiarization phase. This action event was looped and repeated six times. The overall length of each test movie was 93.97 seconds. Infants had to look at the screen during the first approach of the agent up to the moment of divergence[4] in the beginning of each test trial to be included in the final sample.

If infants did not watch the approach of the agent until the moment of divergence, we repeated the same test trial. If they did not look at the approach of the agent until the moment of divergence for the second time either, we proceeded to continue presenting the second test trial. Hence, we started measuring the looking time duration of infants at each test event only after the moment of divergence. If the infant looked away from the screen for two consecutive seconds or more during the test, the software automatically started the next test trial (all the stimulus presentations were infant controlled). The maximum looking time duration was 86.94 sec. If the infants kept looking at the screen for the maximum duration of the test event the software started to show the next test trial.

Following the first pair of test trials, we repeated the whole procedure but with shortened *group induction* and *action familiarization phases* following the protocol of Powell & Spelke (2013). Infants viewed a *group reminder phase* in which only the test agent and the first of the three vertically aligned ingroup agents from the top were dancing together (as illustrated in Figure 4.1e), while the outgroup agent was dancing alone. This group reminder phase was immediately followed by *a second action familiarization phase* where the test agent performed two goal-directed actions. This reminder movie, with no attention-getter in between, took a total of 54.9 seconds. This was followed by the presentation of two test trials.

The second *group reminder phase* involved the test agent and the second ingroup agent from the top dancing together (as can be seen in Figure 4.1f), and the outgroup agent dancing

[4]Even though the "moment of divergence" started 7100 ms into the test-stimulus, due to a scripting error the moment of divergence was defined as 7033 ms in the script for online looking-time calculations. Because this mistake was discovered only after data collection and analysis for Experiment 1, the same value was kept for all the experiments. In principle this 67 ms difference could had affected the trial termination but it never did, hence it is not likely that this difference had affected the results.

alone. The action familiarization phase was exactly the same as described above and was again followed by two test trials.

We counterbalanced the order of efficient and inefficient test movies throughout the procedure across infants. Half of the infants first saw the efficient test movie in the order of AB-BA-AB and the other half first saw the inefficient test movie in the order of BA-AB-BA (i.e. with A standing for efficient action, and B standing for inefficient action).

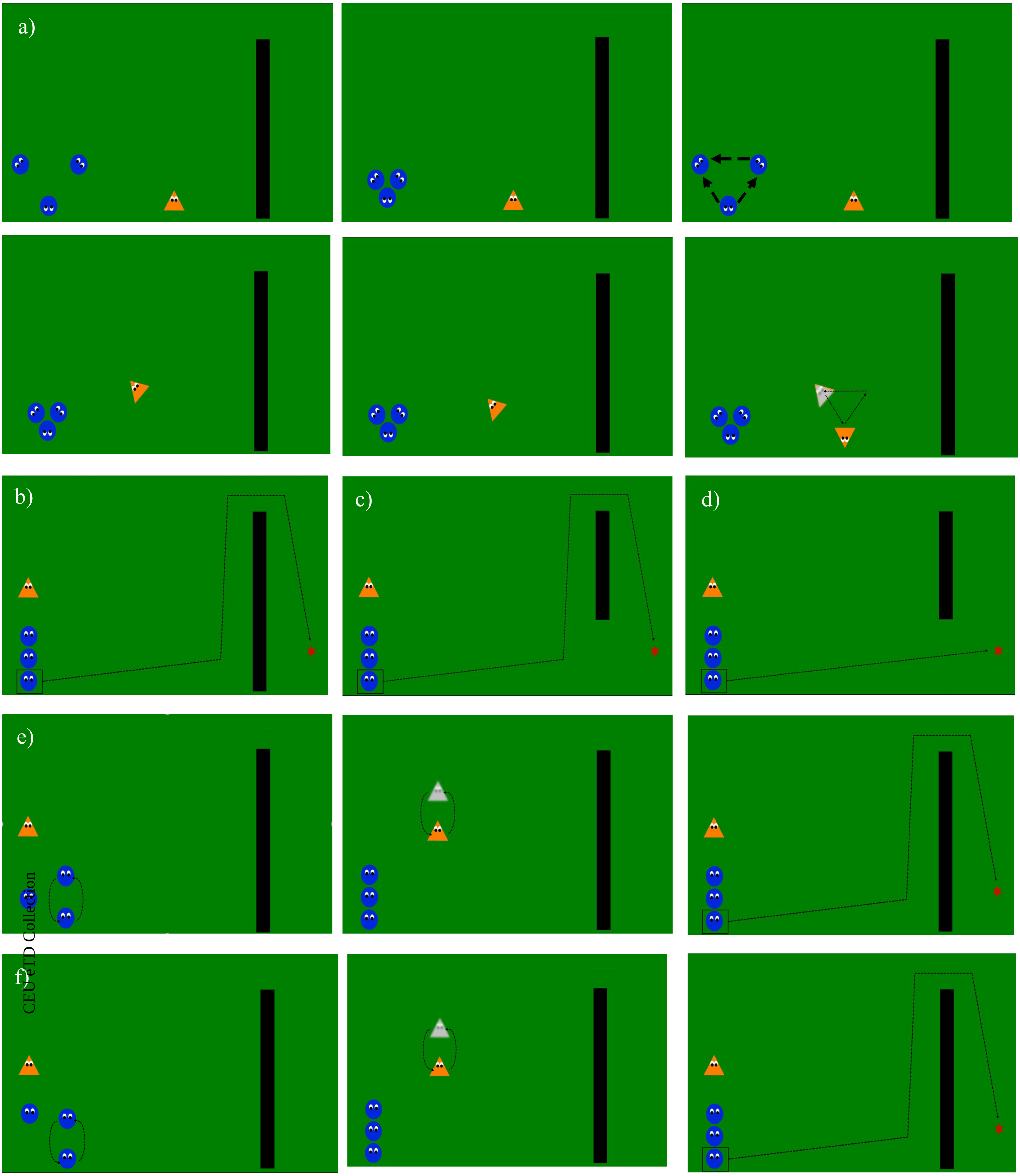
a)
b)
c)
d)
e)
f)

Figure 4.1. Schematic depictions of the stimuli showing a) the dancing by the ingroup agents and by the outgroup agent, b) the action familiarization phase, c) the inefficient action test event, d) the efficient action test event, e) the group reminder and the second action familiarization phase, f) the second group reminder and the third action familiarization phase. Note that the dashed lines represent the trajectory of the actions only schematically: in the actual video stimuli going-above-the-barrier actions looked like jumping over the barrier and the agent did not follow straight lines.

4.2.4 Data exclusion

We excluded all the participants who did not produce reliable looking time data in two consecutive test events. This meant that we excluded infants if their looking time data for the efficient and for the inefficient action came from different test pairs (e.g. test 1 & test 3 or test 4 & test 6). This pair-wise exclusion was chosen in order to preclude the possibility that test trial order could become a factor confounded with test trial type (which could happen by chance, if the valid trials of the two test events were differently distributed across the events in the test phase). Infants also needed to look at least two seconds or longer at each test trial to be included in the final sample, and they should produce maximum looking time in two consecutive test trials.

Furthermore, to be included in the final sample the infants needed to observe:

- At least two dancing actions in *the group induction phase* by two groups of agents (the criterion was one dancing action by each group during the group reminder phases)

- At least two actions during the first *action familiarization phase* (the criterion was at least one action by the ingroup agent during the remaining two action familiarization phases)

- The movement of the agent until the moment of divergence for each test trial

Mothers were instructed to close their eyes and not to interfere with their baby's behavior. If they nevertheless behaved in a way that influenced their infants' looking behavior to the screen, we excluded that participant due to parental interference.

If the baby cried, moved a lot due to fussiness or revealed any sign of discomfort we excluded that infant.

Since our test events were controlled by the infants' own looking behavior (as the software automatically proceeded to show the next test trial as soon as the infant looked away from the screen for two consecutive seconds or more) there were cases when the software proceeded with the presentation of the next trial losing infant's gaze, which resulted in premature termination of the test trial infant was viewing. This premature termination of the test trials was either due to infants' pointing at the screen or their clapping during the test trials as the pointing finger or clapping hands created a shadow on the eye of the infant which interfered with the continuous registration of the data. Additionally, when infants yawned, laughed, coughed or leaned forward towards the screen the software also lost the infants' gaze on the screen even though they were still attending to the test event. We therefore also excluded test trials where such an "eye-tracker registration error" occurred (see Appendix 2 for the trials excluded due to this error).

4.2.5 Results

4.2.5.1 Looking time analyses

Looking time behaviors of six infants who participated in the piloting session with the exact same procedure described above were coded offline, frame-by-frame, from the videos recorded at 25 frames per second. Inter-rater reliability between my offline codings and those of the software was high for test trials (r = .99 for the two different types of test trials pooled together, the mean absolute difference between the two codings was 2.83 s, SD = 4.13 s). Hence, we relied on the looking time data registered by the software for our experimental sample.

We summarize infants' looking time durations to familiarization events in Table 1 of Appendix 2. Infants' looking times to test trials were averaged separately for the two test trial types (efficient vs. inefficient). Next, these averaged looking time data in seconds were log-10 transformed (Csibra et al., 2016). Data were collapsed across participant gender as preliminary analyses showed no significant interaction of this factor with test trials (F (1,22) = .35, p = .56, partial η^2 =. 02). There were 12 infants who saw the efficient action test first, and 12 infants who saw the inefficient action test first. Which test trial was seen first did not interact with the test trials (F (1, 22) = 1.02, p = .32, partial η^2 = .04). There were 12 infants who viewed blue circles as the ingroup agents, while the remaining infants viewed orange triangles as the ingroup agents. Which color-and-shape the ingroup test agent was viewed did not interact with the test trials (F (1, 22) = .45, p = .5, partial η^2 = .02). Half of the infants saw the ingroup agents moving first, while the other half saw the outgroup agent moving first in the group induction phase. Which group of agents moved first in the group induction trials did not interact with test trials F (1, 22) = .05, p = .83, partial η^2 = .002).

There was a statistically significant difference in infants' looking times induced by the efficient versus the inefficient test trials (t (23) = -2.18, p = .04, Cohen's d = .43; a Wilcoxon signed-rank test on raw looking times, Z = -2.11, p = .03). As can be seen in Figure 4.5, infants looked longer at the inefficient action events (M = 33.34 s, SD = 26.22 s) in comparison to the efficient action events (M = 22.85 s, SD = 26.22 s).

4.2.5.2 Pupil size analyses

There was a 500 ms long baseline window starting at 1000 ms until 1500 ms for each test movie during which there was no motion, and the infants viewed the agents all stacked up on the left side of the screen as depicted in Figure 4.1b. Given that we did not have a strong hypothesis about when exactly the predicted effect may take place we used a time window of measurement until one action was completed (i.e. until the agent has completed its action of taking the star back to the left side of the screen from where the action was originally initiated) for the duration of 15500 ms, which directly preceded by the baseline window. The measurement window started at 1500 ms with the agent starting his approach and ended at 17000 ms with the agent and the star on the left side of the screen.

The pupil size data were averaged into 100 ms bins. Baseline pupil size was calculated for each trial by averaging over the five baseline bins (from 1000 ms to 1500 ms). For each bin a percentage change from the baseline was calculated by subtracting the trial's baseline pupil size from the current bin's pupil size and dividing the difference by the trial's baseline pupil size. We further calculated the percentage of change of each baseline corrected data point from the baseline window given that Eyelink reports the pupil size in arbitrary units based on the number of thresholded camera sensor pixels. Outliers that were three standard deviations above or below the mean for each bin were excluded.

For each participant two time-series of pupil size values (one for efficient and one for inefficient test trials) were generated by averaging the bins across trials within each test trial type. These were further averaged into two time-series of group averages and plotted in order to identify the time widow for comparison of pupil dilations elicited by efficient and inefficient test events. For each participant the values within the chosen time-window (11700 ms – 17000 ms) were averaged and the individual averages were compared using both a repeated-measures t-test and its non-parametric equivalent. Note that at 9 seconds the agent already contacted the object, and at 11.7 seconds the agent was seen moving either above (for inefficient test trials) or below (for efficient test trials) the barrier taking the object back to where the agent's initial movement started from.

The results confirmed our hypothesis showing greater pupil dilation at the inefficient action test events in contrast to the efficient action test events (t (20) = -2.36, p = .029, Cohen's *d* = .52; Wilcoxon Signed Ranks Test Z = -1.99, p = .046), see Figure 4.2. Note that we had to exclude four infants due to missing data.

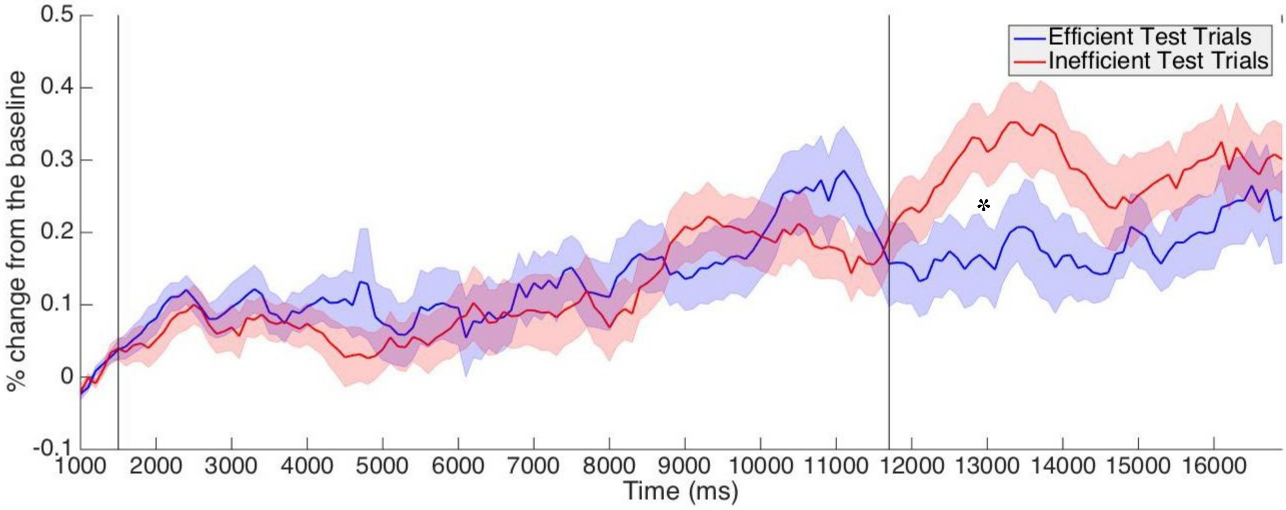

Figure 4.2. Temporal course of percentage of change in pupil size from the baseline (1000-1500 ms) for efficient and inefficient test trials in Experiment 1a. The asterisk marks a statistically significant difference between the time points of 11700-17000 ms.

4.2.6 Discussion

Eleven-month-olds upon viewing an animated character approaching a goal object by going over a barrier in the familiarization phase expected the agent to adapt its course of action to the constraints of the physical environment in the test trials. When the barrier was shorter, allowing the agent to pursue a straight-line approach to the goal, infants expected the agent to follow the efficient straight path and were surprised, as their longer looking time revealed, when the agent repeated its (now inefficient) previous movement of going above the barrier in the test trials. Despite a relatively lengthy group induction phase that infants viewed before the action familiarization scene, our paradigm replicated previous findings which documented that infants expect agents to approach their goals through the most efficient path available (Csibra, Gergely, Bíró, Koós, & Brockband, 1999; Gergely, Nadasdy, Csibra, & Biro, 1995; Biro, Verschoor, & Coenen, 2011; Csibra, 2008a). Our finding of increased pupil size that was induced by observing the inefficient goal approach further indicates that infants payed increased attention to the unexpected inefficient action, suggesting that infants found the inefficient approach cognitively more demanding to accommodate. After all, when the physical change in the length of the barrier allowed the agent to approach the same goal object more efficiently by following an unobstructed direct straight path leading to it during the test trial, the agent was not expected to repeat its previous movement around the barrier to reach the goal object.

We should also point out that the efficient goal approach performed during the action familiarization phase did not correspond to the shortest possible pathway available. The agent could have started to move upwards right away to pass over the barrier and land straight on the object positioned on the bottom right. Instead the agent first moved horizontally towards the goal object but had to change its pathway as it got close to the barrier as if it was only reacting to the presence of the barrier in its path when getting close enough to it. However, this did not prevent infants from attributing a goal to the action they viewed during the familiarization phase or from forming expectations about how the agent should approach the goal object when there was a relevant change in the environmental constraints during the test phase. It appears that infants can form efficiency expectations rather flexibly. However, this does not mean that infants will not be sensitive to the differences between the two alternative efficient action trajectories perceived, especially when different agents pursue slightly different but – given the circumstances –sufficiently efficient trajectories toward the goal.

In fact, in our further experiments we shall present infants with familiarization scenes in which different agents will perform their efficient goal approach starting from different initial positions. In other words, infants in these experiments will be familiarized to efficient actions that follow slightly different trajectories to the goal and will be presented with another efficient trajectory performed by another agent during test. This requires us to first establish whether observing different efficient action trajectories during the action familiarization phase would be a factor preventing them from forming predictions about the future behaviors of the agent in the changed environmental constraints of the test situation. Hence Experiment 1b aims to replicate the findings of Experiment 1a using a slightly different scenario that present the alternative efficient approach trajectories taken by different agents in Experiments 3 and 4.

4.3 Experiment 1b. Individual efficiency: efficient approach trajectories

In the planned series of experiments (see Table 4.1) the action trajectories that each individual agent would follow during the efficient action familiarization phases of Experiment 3 and 4 are slightly different from each other due to the spatial arrangement of the agents on the scene. For example, as can be seen in Figure 4.3.a, in Experiment 3 (ingroup efficiency) infants will observe three ingroup agents vertically stacked one above the other on the left side of the screen. Due to the agents' different original position on the scene from where they start their respective actions, their action trajectories during the efficient approach familiarization events will be slightly different until they reach the moment of divergence. Moreover, unlike in Experiment 1a, in the further series of experiments the agent acting during the test phase will be a different agent than the one acting during the action familiarization events. Hence at the test phase infants will not be viewing the same action trajectories they saw during the action familiarization phase.

Given these differences, Experiment 1b aimed to first test whether familiarizing infants with slightly different efficient approaches (as in Experiments 3 and 4, see Figure 4.3) and showing them a yet again different efficient trajectory during the test event would not prevent them from forming similar teleological inferences to guide their expectations about the agents' future actions. Hence Experiment 1b was designed to test for infants' sensitivity to the different movement trajectories of the goal-directed actions that will be presented to them during the efficient approach scenes we planned to use in Experiment 3 (dubbed as "low-trajectory") and Experiment 4 (dubbed as "high-trajectory").

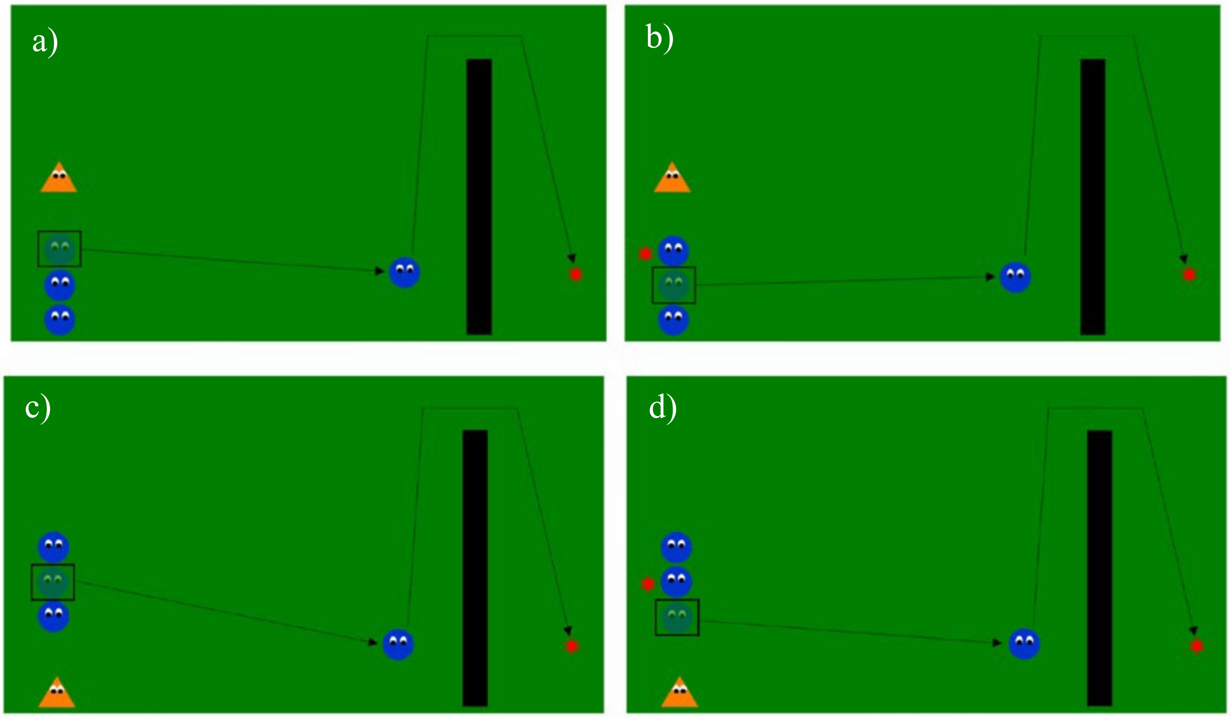

Figure 4.3. Schematic depiction of the *low-trajectory and high-trajectory*: efficient actions to approach the goal object by the two agents during the action familiarization phases in Experiment 3 ("ingroup efficiency") (a-b) and in Experiment 4 ("outgroup efficiency") (c-d). Note the slight variations in the action trajectories of each agent.

4.3.1 Participants

Participants were 24 infants (12 females). Mean age of the participants was 336.75 days (11 months) (range 324-349 days; 10;20 to 11;12, *SD* = 7.62 days). An additional 28 infants were tested but excluded from the final sample due to inattentiveness (n = 1), fussiness (n= 3), looking-times of less than 2 seconds per trial (n = 9), premature termination of a trial resulting from a lack of eye-tracking data for 2 consecutive seconds rather than looking away from the screen ($n = 9$), software error (i.e. script crashing in the middle of the experiment, $n = 4$), and experimenter error (i.e. calibration error, $n = 2$).

Half of the infants viewed "low trajectory" actions, and the remaining half viewed "high trajectory" actions during action familiarizations.

4.3.2 Procedure and stimuli

The procedure was the same as in Experiment 1a except that: 1) There was only one agent present in the display (either an orange triangle or a blue circle, counterbalanced across participants), 2) There was neither a group-induction phase nor a group-reminder phase presented. During all the three *action familiarization phases,* the agent efficiently approached the goal object (four times in the first action familiarization phase, and twice during the other two action familiarization phases) but with either through low-action trajectories or high-action trajectories. The low-action trajectories were the same as the action trajectories through which the two ingroup members will approach the goal object in Experiment 3. The high-action trajectories corresponded exactly to the action trajectories through which the two ingroup members will approach the goal object in Experiment 4. See Figure 4.4 for the depiction of the events in the "*low-trajectory*" and in the "*high-trajectory*".

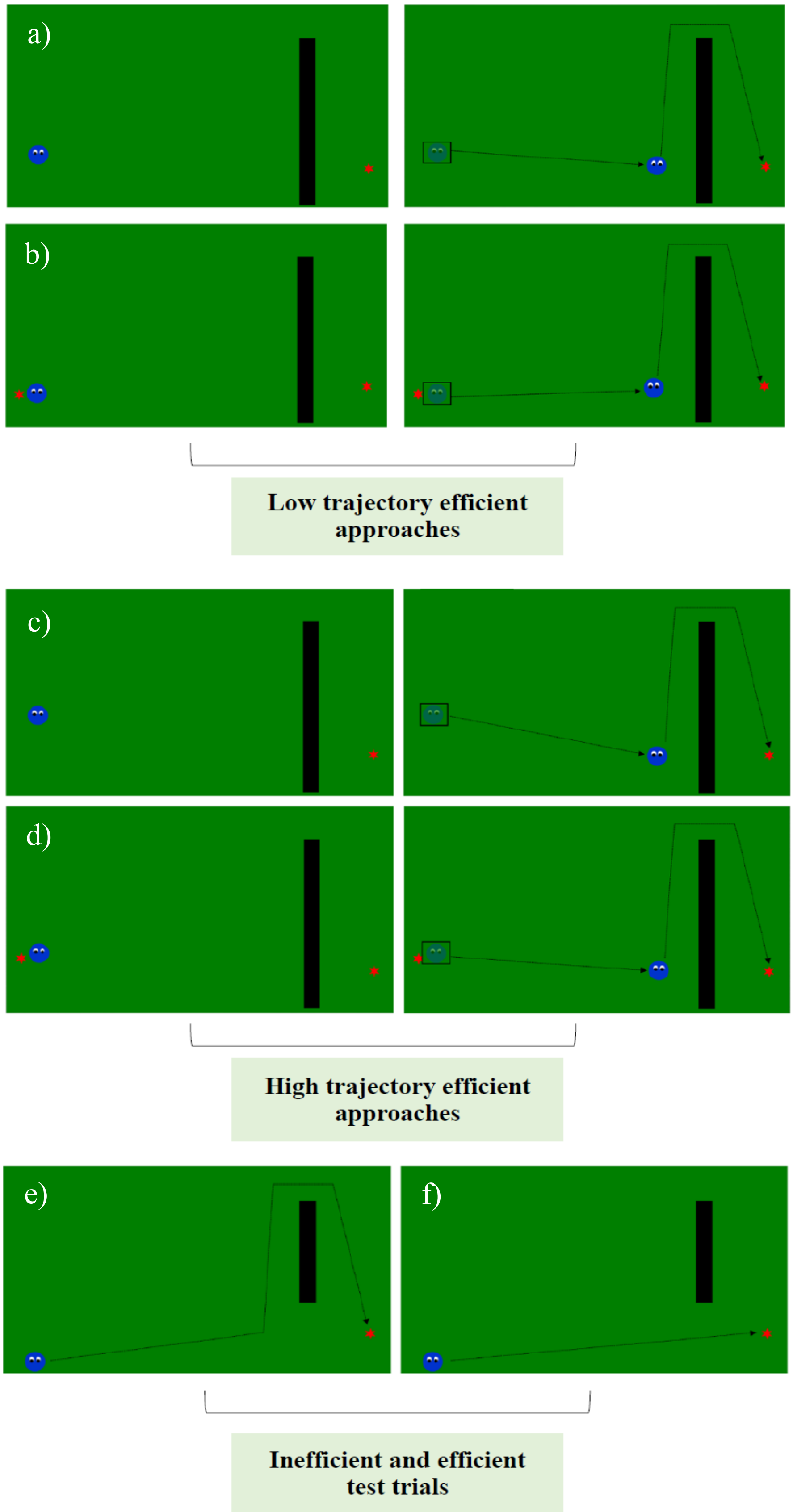
a)
b)
Low trajectory efficient approaches
c)
d)
High trajectory efficient approaches
e)
f)
Inefficient and efficient test trials

Figure 4.4. Schematic depiction of the stimuli showing: (a-b) low trajectory efficient approaches of the single agent towards the goal object, (c-d) high trajectory efficient approaches of the single agent towards the goal object during the action familiarization phase, (e-f) the inefficient and efficient test trials.

The test phase was exactly the same as in Experiment 1a, except that there was a single agent on the scene as in the familiarization phase. The barrier that was viewed in *the action familiarization phase* was shortened by 40% of its size (leaving an unobstructed direct approach route to the goal object, see Figure 4.4e, 4.4.f). The test agent who positioned in the bottom of the screen approached the target object either efficiently or inefficiently in the test trials.

Infants viewed action familiarization movies in the following order, yielding six test trials in total:

- Action familiarization phase (four action events)
- Test phase (two test trials)
- Second action familiarization phase (two action events)
- Test phase (two test trials)
- Third action familiarization phase (two action events)
- Test phase (two test trials)

4.3.3 Results

4.3.3.1 Looking time analyses

As in Experiment 1a, infant's looking times to the test trials were averaged separately for the two test-trial types (efficient vs. inefficient). Next, these averaged looking-time data in seconds were log-10 transformed (Csibra et al., 2016). Data were collapsed across participant gender as preliminary analyses showed no significant interaction between this factor and the test trials ($F\,(1,22) = .89$, $p = .36$, partial $\eta^2 =. 04$). There were 12 infants who saw the efficient action test first, and there were 12 infants who saw the inefficient action test first. Which test trial was seen first did not interact with test trials (*F* (1, 22) = .04, *p* = .84, partial η^2 = .002). There were 12 infants who viewed a blue circle being the agent, while for the remaining participants the agent was an orange triangle. The different color and shape of the agent observed did not interact with the test trials (*F* (1, 22) = .203, *p* = .66, partial η^2 = .009). Half of the infants viewed only low trajectory events in the action familiarization phase whereas the other half viewed only high trajectory events. The trajectory of the events did not interact the with test trials (*F* (1, 22) = .52, *p* = .47, partial η^2 = .02).

There was a statistically significant difference in infants' looking time to efficient and inefficient action test trials (*t* (23) = -2.25, *p* = .03, Cohen's *d* = .48; a Wilcoxon signed-rank test on raw looking times, *Z* = -2.09, *p* = .037). As can be seen in Figure 4.5, infants looked longer to inefficient test trials (*M* = 31.19 s, *SD* = 19.38 s) than to efficient action test trials (*M* = 23.29 s, *SD* = 17.59 s).

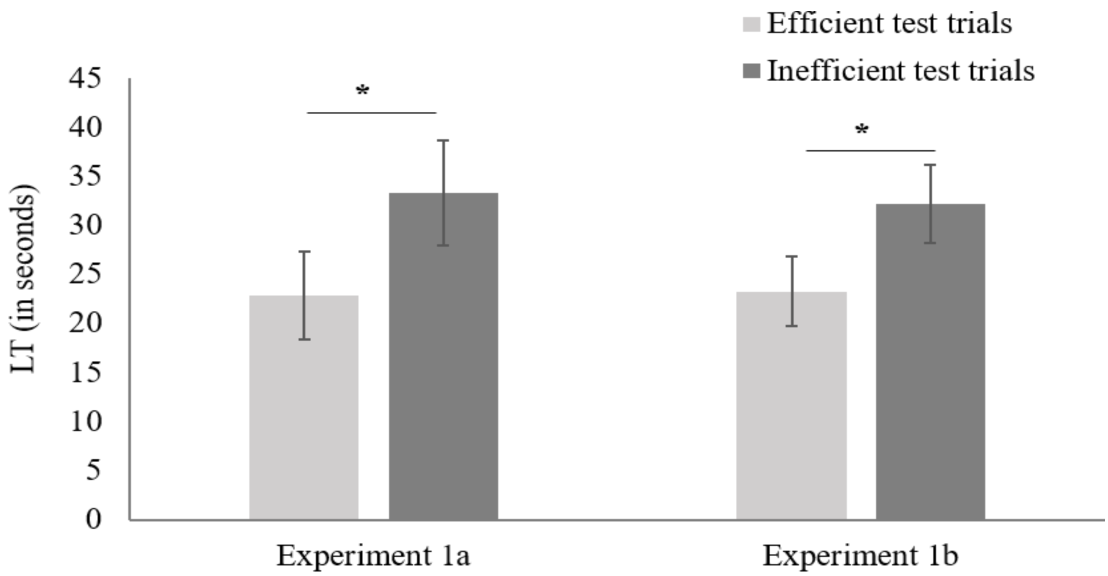

Figure 4.5. Mean looking times in seconds at the two types of test trials for infants in Experiments 1a and 1b. Error bars show ±1 *SE*. Asteriks mark a statistically significant difference for the log-10-transformed data.

4.3.3.2 Pupil size analyses

For each participant two time-series of pupil-size values (one for efficient and one for inefficient test trials) were generated by averaging the bins across trials within each test trial type exactly as explained in detail for Experiment 1a. For each participant the values within the time window chosen in the previous experiment (11700 ms – 17000 ms) were averaged and the individual averages were compared using both a repeated-measures t-test and its non-parametric equivalent. As depicted in Figure 4.6, there was a greater pupil dilation response induced by the inefficient action test events than by the efficient action test events (t (22) = -3.04, p = .006, Cohen's d = .64; Wilcoxon Signed Ranks Test Z = -2.95, p = .002). Note that one infant had to be excluded from the analysis due to missing data.

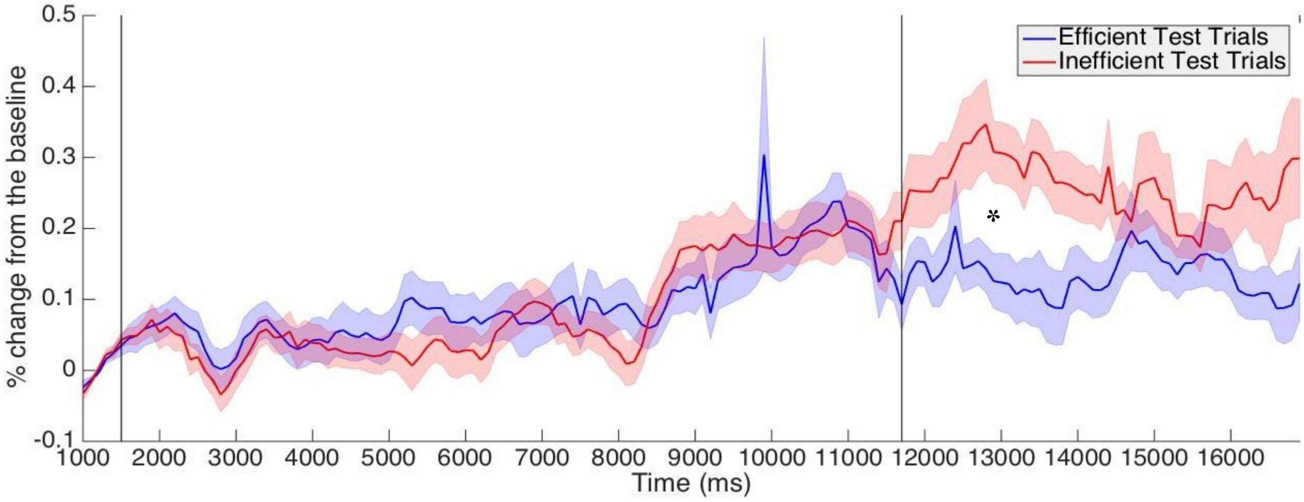

Figure 4.6. Temporal course of percentage of change in pupil size from the baseline for efficient and inefficient test trials in Experiment 1b. The asterisk marks a statistically significant difference between the time points of 11700 ms -17000 ms.

4.3.4 Discussion

Despite being familiarized with slightly different efficient approaches (exactly as they will view in Experiments 3 and 4) infants were able to make use of this information to represent a similar action trajectory they viewed in the test as being inefficient given the change in the length of the barrier during the test phase. As in Experiment 1a, 11-month-old infants again looked longer at the inefficient action tests in comparison to the efficient action tests. Moreover, we again found a significant increase in pupil size in response to inefficient actions. These findings show that different trajectories in the efficient approach actions did not prevent infants from forming teleological inferences about the expected future actions of the agent. This was so even when the agent was starting his action from a higher initial position in the display (see Figure 4.4c), which made its action trajectory deviate from the shortest path to the goal to a greater extent than what was observed in the familiarization scenes of Experiment 1a. Despite these differences in the efficient approach trajectories during the familiarizations, infants still expected the agent to approach the goal object by pursuing a straight-line trajectory when the barrier was made shorter allowing the agent to access the object directly. Infants looked longer when the agent moved similarly as it did during the familiarization, by going over the barrier to descend on the other side in order to access the goal object.

However, this finding in-and-of itself does not suggest that infants are not sensitive to variations in the efficient action trajectories observed. Previous research showed that infants can take the variability in the performed goal-directed actions into account and guide their reasoning about agentive properties of observed entities based on whether they are able to choose their action trajectories freely (Biro & Leslie, 2007; Csibra, 2008a; Luo, 2011; Shimizu & Johnson, 2004). Equifinal variations of the action thus seems to provide an agency cue for young infants inducing them to represent even unfamiliar entities' actions as goal directed. In Experiment 1b we simply aimed to show that these slightly different efficient trajectories would not prevent infants from forming teleological inferences about expectable future actions of the agents. The present findings indeed suggest that infants do not base their inferences about the agent's expectable (and different) future approach to the goal in changed circumstances by taking into consideration of the small differences in their efficient trajectories during familiarization. Rather, they seem able to form robust predictions guided by the principle of rationality in interpreting the variable actions they have observed. However, as outlined in the introduction of this chapter, we do not yet know whether this robust expectation of efficiency could be modulated by viewing agents acting alike in social groups.

4.4 Experiment 2. Ingroup inefficiency

Experiment 2 aims to directly pit against each other the relative effects of shared movement repertoires of agents belonging to the same social group on the one hand, and their efficiency of goal-approach, on the other. Previous research documented that infants expect common behaviors to be performed by members of the same social group (Powell & Spelke, 2013). However, we do not yet know whether this expectation for shared movement repertoires can trump infants' robust expectation for efficiency of goal approach. In this experiment, infants viewed the group induction scene with four agents exactly as it was in Experiment 1a (three agents were in the same social group [ingroup agents] and the fourth one was not [the outgroup agent]). Later they were familiarized to shared movement repertoires of ingroup agents: two ingroup agents both approached the goal object by moving above a barrier to reach the goal object on the other side, alternately. Critically the barrier this time was already short during action familiarization phase, unlike in the previous experiments, and did not block direct access to reach the goal object. Still, the ingroup agents performed detoured movements in obtaining the goal object during the action familiarization phase. At test, infants watched the third ingroup agent either performing the same detour movement over a short barrier just like the other group members did during the familiarization phase (inefficient trial) or a taking a direct path to the goal (efficient trial).

If infants represent the detour movements of the two ingroup agents in the familiarization phase as constituting the shared movement repertoire of the group, as we predicted, then they should expect the third ingroup member to also act like them during the test phase. This predicts longer looking times when seeing the third group member perform a dissimilar (even though efficient) goal approach than when observing a similar (but inefficient) goal approach. On the other hand, if infants expect an efficient goal approach from the third ingroup member during the test, independently of how other group members moved before, then they should look longer at the inefficient-but-similar action trials than at the efficient-but-dissimilar action trials. We further predicted that relative pupil dilation should reflect the pattern of looking times induced during the test events (showing greater pupil dilation for efficient action this time, in line with our prediction).

4.4.1 Participants

Participants were 24 infants (15 females). Mean age of the participants was days 332.8 (10 months 28 days) (range 319-348 days; 10;15 to 11;13, *SD* = 9.13 days). An additional 28 infants were tested but excluded from the final sample due to fussiness (*n*= 5), looking-times of less than 2 seconds per trial (*n* = 11), premature termination of a trial resulting from a lack of eye-tracking data for 2 consecutive seconds rather than looking away from the screen (*n* = 12).

4.4.2 Procedure and stimuli

The procedure was similar to the one used in Experiment 1a, except that during the *action familiarization phase* two different ingroup agents approached the goal object twice (rather than one of the ingroup agents approaching the goal four times as in Experiment 1a). During the second action familiarization phase, it was the first ingroup agent on the top acting twice; and during the third action familiarization phase, it was the second ingroup agent from the top acting twice. Also, unlike in Experiment 1a, the barrier between the agents and the goal was short during the action familiarization phases (i.e. it did not block a direct straight-line approach route leading to the goal). This allowed the agents to follow an efficient direct approach route to reach the goal object already during the familiarization. However, despite the availability of the unblocked direct straight-line pathway to the target object, both agents repeatedly approached the object by making an unnecessary detour over the barrier, resulting in an inefficient approach. Please see Figure 4.7 for the depiction of the events.

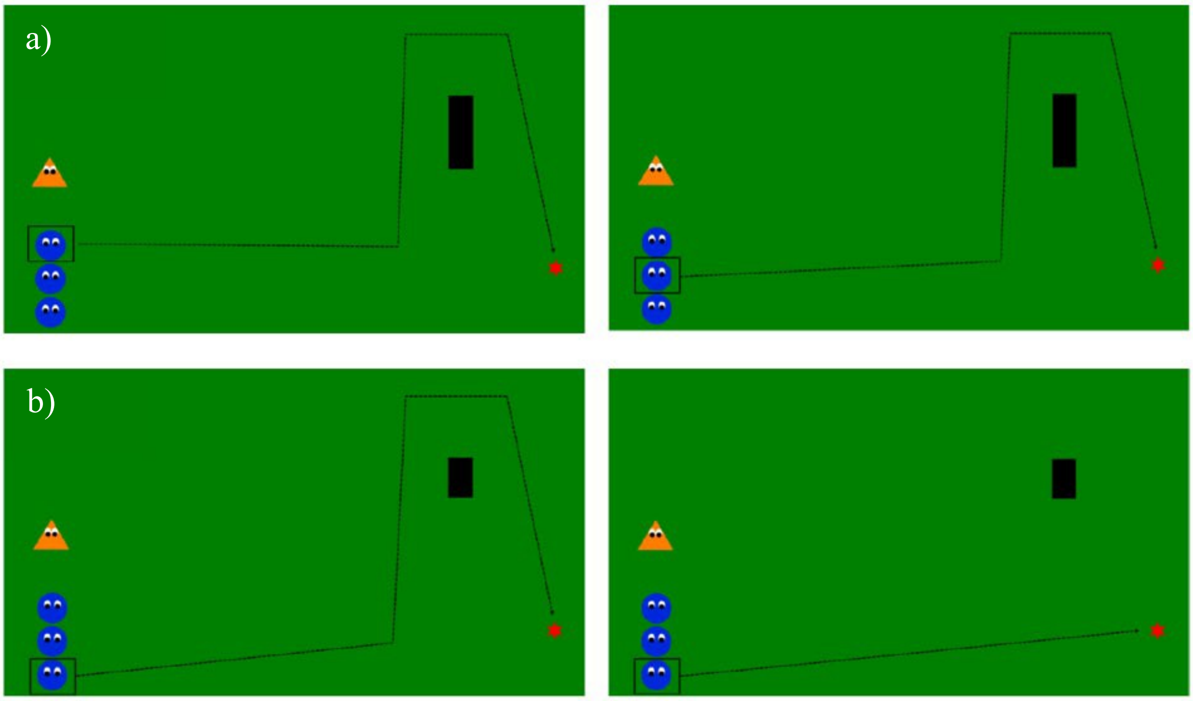

Figure 4.7. Schematic depiction of the stimuli showing a) inefficient actions performed by the two ingroup agents during the action familiarization phase, b) inefficient and efficient test events, respectively.

The test phase was similar to that used in Experiment 1a. The barrier that was viewed in *the action familiarization phase* was again shortened by 40% of its size as can be seen in Figure 4.7.b, and the test agent positioned in the bottom of the screen, whom the infants did not see acting during the action familiarization, approached the target object either efficiently or inefficiently in the test trials.

4.4.3 Data exclusion

Data were excluded based on the criteria outlined for Experiment 1a.

4.4.4 Results

4.4.4.1 Looking time analyses

Looking time data were log-10 transformed as in the previous experiments. Data were collapsed across participant gender as preliminary analyses showed no significant interaction of this factor with test trials (F (1,22) = 1.43, p = .24, partial η^2 =. 06). There were 12 infants who saw the efficient action test first, and the remaining 12 infants saw the inefficient action test first. Which test trial was seen first did not interact with test trials (F (1, 22) = 2.38, p = .14, partial η^2 = .1). There were 12 infants who viewed blue circles as the ingroup agents, while the remaining participants viewed orange triangles as the ingroup agents. Which color- and shape- the ingroup agents were viewed significantly interacted with test trials (F (1, 22) = 7.1, p = .01, partial η^2 = .24). Infants viewing the blue circles as the ingroup agents looked longer to inefficient action test trials (M = 37.78, SD = 20.72) in comparison with the efficient test trials (M = 29.34, SD = 27.3). This difference was statistically significant with a parametric test on log transformed data, (t (11) = -2.59, p = .02, Cohen's d = .74), but not with a non-parametric test on raw looking times (a Wilcoxon signed-rank test, Z = -1.88, p = .06). Infants viewing the orange triangles as the ingroup agents looked longer to efficient action test trials (M = 23.52 s, SD = 16.88 s) in comparison with the inefficient test trials (M = 17.06 s, SD = 14.76 s). This difference was not statistically significant (t (11) = 1.27, p = .23, Cohen's d = .36; a Wilcoxon signed-rank test on raw looking times, Z = -1.09, p = .3). Half of the infants saw the ingroup agents moving first, while the other half saw the outgroup agent moving first in the group induction. Which group of agents moved first in the group induction trials did not interact with test trials F (1, 22) = .26, p = .61, partial η^2 = .01).

There was no statistically significant difference in infants' looking time to efficient and inefficient action test trials (t (23) = -.68, p = .5, Cohen's d = .14; a Wilcoxon signed-rank test on raw looking times, Z = -.57, p = .58). As can be seen in Figure 4.13, infants looked equally long to inefficient (M = 27.42 s, SD = 20.53 s) and efficient action test trials (M = 26.43 s, SD = 22.4 s).

4.4.4.2 Pupil size analyses

For each participant the values within time window chosen in the previous experiments (11700 ms – 17000 ms) were averaged and the individual averages were compared using both a repeated-measures t-test and its non-parametric equivalent. As depicted in Figure 4.8 there was no differential pupillary response at two test events (t (20) = -.74, p = .47, Cohen's d = .16; Wilcoxon Signed Ranks Test Z = -.99, p = .34). Note that we had to exclude three infants due to missing data.

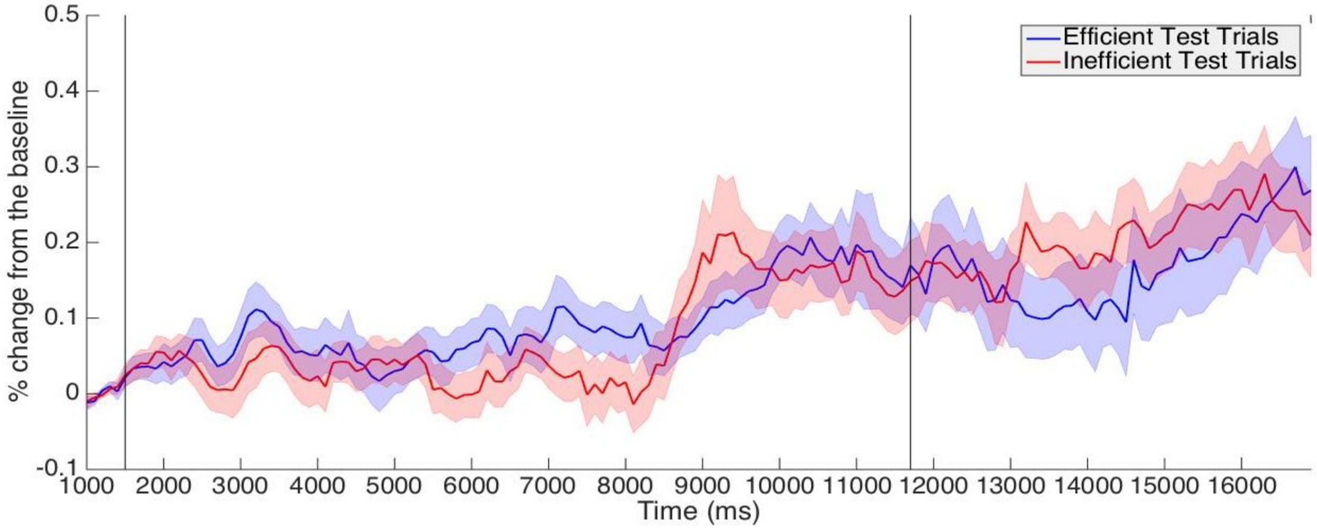

Figure 4.8. Temporal course of percentage of change in pupil size from the baseline for efficient and inefficient test trials in Experiment 2.

4.4.5 Discussion

We predicted infants to represent the detour movements of the two ingroup agents in the familiarization as belonging to the shared movement repertoire of the group and guide their inferences about the third ingroup agent's behavior accordingly. However, infants did not look longer at the efficient test trials than at the inefficient test trials. Surprisingly, neither did they expect efficient approach from the third ingroup member. Pupillary analysis also failed to detect a difference between infants' responses to the two test events.

Looking time findings revealed that infants viewing the ingroup agents as blue circles looked longer at inefficient trials in contrast to efficient trials. For infants viewing the ingroup agents as orange triangles the difference between the looking times at the two test trials was not statistically significant. We did not predict this interaction, and the previous experiments did not reveal such an interaction between the appearance of the agent and the test trials. The appearance of the ingroup and outgroup agents was a factor that we counterbalanced within each experiment, and we admittedly do not have a good account to explain this interaction. Rather the null findings for the general pattern of looking times revealed for efficient and inefficient trials suggest that 11-month-old infants 1) do not expect a novel ingroup member to perform the inefficient actions that its group members had previously performed and 2) they do not expect a novel ingroup member that they did not see acting before to pursue an efficient approach action towards a goal candidate. Below, we discuss possible reasons for this.

Considering that Powell and Spelke (2013) familiarized infants with jumping and sliding motions, the inefficient actions infants in Experiment 2 viewed during the familiarization phase were also likely to generalize to a third agent as long as it belonged to the same group of the agents acting during the familiarizations. After all, inefficient approach actions performed by group members could have signaled the infants that the actions might have had social relevance over and beyond its instrumental function, as the movement trajectory made it transparent to the infants that if the agents had the goal of retrieving the object only, they could had done so by moving under the barrier in a straight motion. Yet the cues of inefficiency and shared movement pattern between group members did not grant the infants with an interpretation that this is how this group of agents act like. It is possible that pre-verbal infants do not represent deviation from the efficiency as a sub-goal, like adults can (Baker, Saxe, & Tenanbaum, 2009). We will elaborate on this point in the general discussion further.

However, infants in our study neither did expect the third group member whom they did not see perform an act during the familiarization phase to approach the goal object efficiently during the test. Is it possible that teleological inferences infants can form are bound to particular agents whom they have seen acting before? We cannot provide an answer to this question based on the findings of Experiment 2 simply because the inefficient approach actions infants were familiarized to in this experiment could have been a reason why they had not attributed a goal to the inefficient behaviors of the agents observed (which would have explained their consequent failure to generate action predictions in the presence of a goal object during the test phase) . However, if that were the case, then infants despite viewing efficient approach actions by two group members during the familiarization phase would not be expected to form any specific expectations regarding the efficiency of the third group member during the test phase either. Experiment 3 aims to test this.

4.5 Experiment 3. Ingroup efficiency

Previous studies examining the teleological interpretation system that infants were assumed to possess only tested infants' expectation of efficiency by showing them an agent acting in a goal directed fashion during a familiarization (or a habituation) phase and presented to infants the same agent acting efficiently or inefficiently in the test phase. Experiment 3 aims to investigate whether infants can form predictions about the possible action route of a novel agent they did not see acting before, after having been familiarized with efficient goal directed actions by its group members approaching the goal object earlier.

4.5.1 Participants

Participants were 24 infants (6 females). Mean age of the participants was 338.17 days (11 months) (range 323-350 days; 10;19 to 11;16, *SD* = 8.59). An additional 19 infants were tested but excluded from the final sample due to experimenter error (*n* = 2), mother interference (*n* = 2), fussiness (*n*= 3), looking-times of less than 2 seconds per trial (*n* = 5), premature termination of a trial resulting from a lack of eye-tracking data for 2 consecutive seconds rather than looking away from the screen (*n* = 7).

4.5.2 Procedure and stimuli

The procedure was the same as in Experiment 2, except that during all the three *action familiarization phases* two different ingroup agents approached the goal object *efficiently* twice (rather than two ingroup agents approaching the target object *inefficiently* twice, as in Experiment 2). See Figure 4.9 for the illustration of the stimulus events. During the test, the long barrier that was present during *the action familiarization phase* was shortened by 40% of its size (allowing for an agent to pass through under it) (see Figure 4.9b), and the test agent positioned in the bottom of the screen, whom the infants did not see acting during the action familiarization phase, approached the target object either efficiently or inefficiently during the test trials.

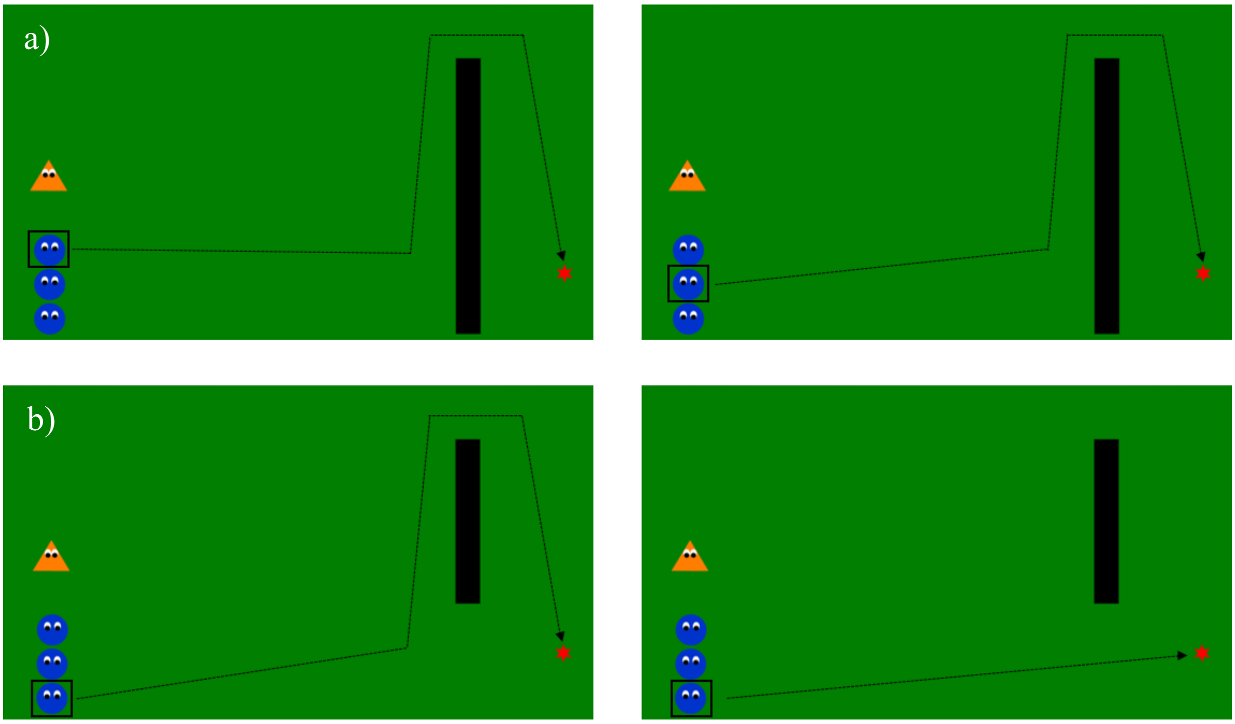

Figure 4.9. Schematic depiction of the stimuli used in Experiment 2 showing: a) efficient actions to the goal taken by the two ingroup agents during the action familiarization phase, and b) inefficient and efficient test events, respectively.

4.5.3 Results

4.5.3.1 Looking time analyses

Looking time data were log-10 transformed exactly as it was in the previous experiments. The data were collapsed across participant gender as preliminary analyses showed no significant interaction of this factor with test trials (F (1,22) = 3.6, p = .07, partial η^2 =. 14). There were 12 infants who saw the efficient test action first, and 12 infants who saw the inefficient test action first. Which test trial was seen first did not interact with test trials (F (1, 22) = .13, p = .72, partial η^2 = .006). There were 12 infants who viewed the ingroup agents as blue circles, while the remaining viewed the ingroup agents as orange triangles. Which color and shape had the ingroup agents did not interact with test trials (F (1, 22) = 1.12, p = .3, partial η^2 = .05). Half of the infants saw the ingroup agents moving first, while the other half saw the outgroup agent moving first in the group induction. Which group of agents moved first in the group induction trials did not interact with test trials F (1, 22) = .05, p = .82, partial η^2 = .002).

There was no statistically significant difference in infants' looking time to efficient and inefficient action test trials (t (23) = -.68, p = .5, Cohen's d = .14; a Wilcoxon signed-rank test on raw looking times, Z = -.37, p = .73). As can be seen in Figure 4.13, infants looked equally long to inefficient (M = 26.32 s, SD = 24.24 s) and efficient action test trials (M = 22.82 s, SD = 22.39 s).

4.5.3.2 Pupil size analyses

For each participant the values within time window chosen in the previous experiments (11700 ms – 17000 ms) were averaged and the individual averages were compared using both a repeated-measures t-test and its non-parametric equivalent. As depicted in Figure 4.10 there was a differential pupillary response induced by the two test events (t (19) = -2.16, p = .04, Cohen's d = .49; Wilcoxon Signed Ranks Test Z = -2.5, p = .01), revealing greater pupil dilation at the inefficient test events. Note that we had to exclude four infants due to missing data.

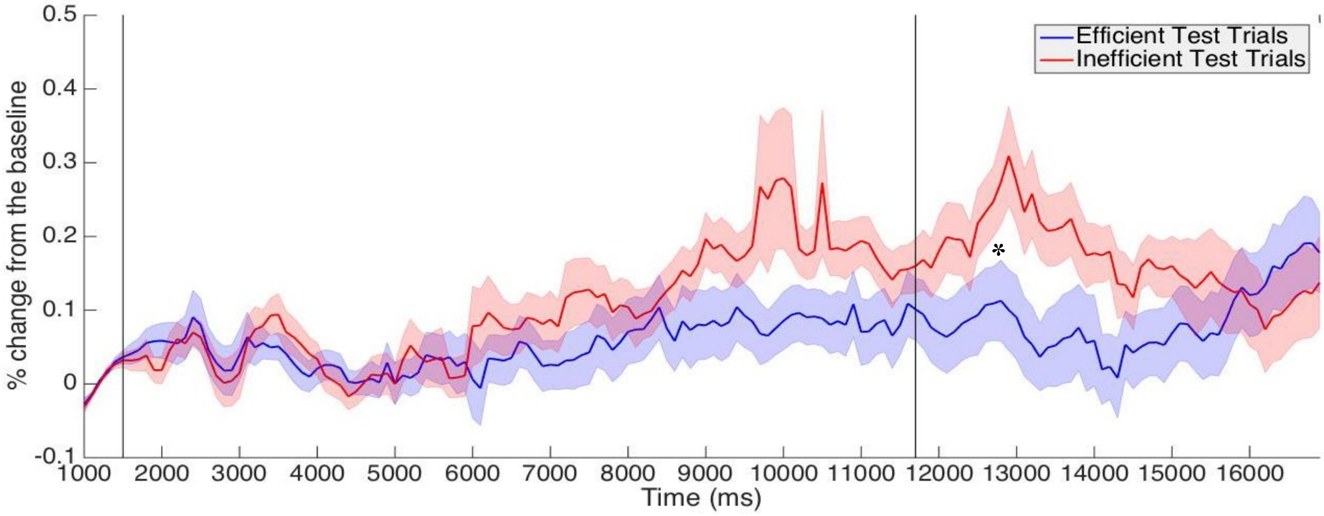

Figure 4.10. Temporal course of percentage of change in pupil size from the baseline for efficient and inefficient test trials in Experiment 3. The asterisk marks a statistically significant difference between the time points of 11700 ms -17000 ms.

4.5.3.3 Pupil size analyses for action familiarizations of Experiment 2 and 3

Despite not finding any evidence for differential pupil dilation in Experiment 2 to efficient and inefficient test trials, it is a possibility that infants' increased pupil response to inefficient trials in Experiments 1a, 1b and 3 might simply stem from the upwards motion involved in these test trials, not genuinely reflecting their expectation of efficient approach. Therefore, we ran a post-hoc comparison of infants' pupil size to efficient and inefficient action familiarization events in Experiment 2 and 3. Note that in these familiarization events, it was only the size of the barrier that was changed as both actions involved the same (upward) motion while the differential length of the barrier rendered the same action either inefficient (Experiment 2) or efficient (Experiment 3).

We analyzed the pupil size data induced by the action familiarization phases of Experiment 2 and 3 exactly as we did for the test trials. There was a 500 ms long baseline starting from 23400 ms to 23900 ms for the second and the third action familiarization scenes during which there was no motion and the agents were all vertically stacked up on the left side of the screen. We excluded the data coming from the first action familiarization phase given that the only 500 ms period in which there was no motion on the scene was directly preceded by the attention getter.

We had the measurement time window until one action was completed for the duration of 15500 ms (from 23900 ms to 39400 ms). This time window was directly preceded by the baseline window. As it was the case for the test trials, the measurement window started at 23900 ms with the agent starting his approach and ended at 39400 ms (when the agent has completed its action of taking the star to the left side of the screen). For each participant two time-series of pupil-size values (one for the action familiarizations of Experiment 2 and one for the action familiarizations of Experiment 3) were generated by averaging the bins across familiarization trials within each experiment.

For each participant the values within the same period chosen for the test trials (which corresponded to the time window between 34100 ms and 39400 ms in the second and the third action familiarization videos) were averaged and the individual averages were compared using both a repeated-measures t-test and its non-parametric equivalent. As depicted in Figure 4.11 we found a differential pupillary response for the two familiarization events. There was a greater pupil dilation for infants watching the efficient familiarization events (t (17) = 2.15, p = .046, Cohen's d = .5; Wilcoxon signed rank tests, Z = -2.11, p = .03) in comparison to those infants watching the inefficient familiarization events. Note that we had to exclude six infants due to missing data.

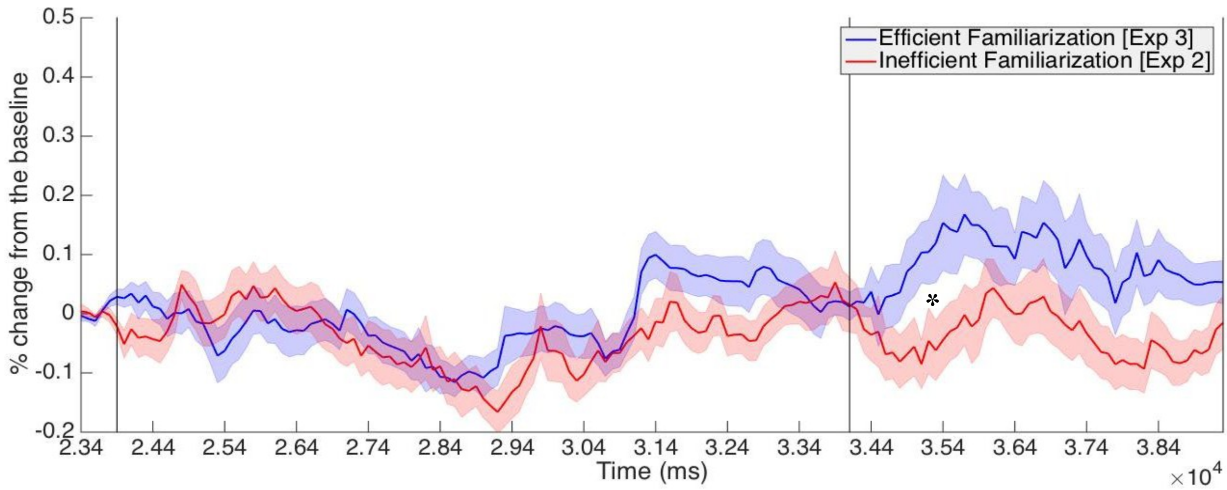

Figure 4.11. Temporal course of percentage of change in pupil size from the baseline for efficient and inefficient actions viewed during the familiarization of Experiment 3 and Experiment 2, respectively. The asterisk marks a statistically significant difference between the time points of 34100 ms - 39400 ms.

4.5.4 Discussion

In Experiment 3 “ingroup efficiency”, infants’ looking times at the two types of test actions (i.e., the efficient test events versus the inefficient test events) were not different from each other. Pupillary data, however, in contrast to the looking time findings documented infants’ increased attention to inefficient trials in contrast to efficient trials as it was found in Experiments 1a and 1b. Given the temporal sensitivity of the pupil size measure, relative pupil dilation was perhaps a more responsive measure of infants’ attention, as it revealed that infants expected efficient approach from the third group member during the test phase. Yet we also suspected that infants’ pupillary response may have been modulated by the upwards motion of the inefficient test trials and so decided to run a post-hoc comparison of infants’ pupillary response to action familiarization events in Experiments 2 and 3. This analysis revealed a greater pupil dilation in response to efficient familiarization events in contrast to inefficient familiarization events.

Note that the findings of a post-hoc pupillary analysis for the efficient and inefficient action familiarization events of Experiment 2 and 3 suggest that infants’ pupillary response might not be merely modulated by the upwards motion of the inefficient action trials during the test (as the action familiarization events in both experiments were the same, both involving upward motion). Rather it appears that pupil dilation in response to the efficient action events that infants observed during the familiarization phase reflects the cognitive effort they engaged in while attributing goals to the actions of agents. This is in line with previous evidence showing that goal attribution is an inferential process and hence requires cognitive effort (Csibra & Gergely, 2007). We therefore suggest that pupil dilation in response to efficient action events during the familiarization phase indicates cognitive processing associated with goal attribution. Pupil dilation in response to inefficient test trials in Experiment 3 on the other hand, might rather reflect the violation of the goal attribution infants generated during the efficient action familiarization events.

However, it is still surprising that this expectation for efficient action during the test phase in Experiment 3 was not evident in the cumulative looking time measure. We consider two possible reasons for the null finding between the looking times induced by the two test events. First of all, 11-month-old infants might not generate expectations about the efficiency of an agent's possible action path if it is the first time that they are observing that agent act towards the goal object. However, this might not be likely given that the pupil measure reflected infants' increased attention to the event they found unexpected (unlike it was the case for Experiment 2). Secondly, 11-month-old infants might be actually generating expectations about the efficiency of a novel agent's possible action path, but this could be in conflict with their expectation that group members should act alike. After all, infants viewed the two same-group agents during the familiarization events moving above the barrier in attaining the target object. The means actions they viewed during this event were efficient due to the presence of the long barrier, which allowed them to attribute a goal to these actions. Given the success of goal attribution, infants might have been able to encode the manner in which the goal was achieved and expected this manner to be copied by the third group member. At the same time, they might have also expected efficient approach action from the third group member given the physical change in the environment. Experiment 4 aims to address this conflict.

4.6 Experiment 4. Outgroup efficiency

In Experiment 4 infants were familiarized to two ingroup agents who went over the top of a long barrier to descend on the other side to reach the goal object as in Experiment 3 (efficient approaches). Yet we tested infants' expectation of a future action path of an outgroup agent in a scene where the barrier was shortened allowing the outgroup agent to either approach the goal object following a straight path (efficient approach) or approach it through the exact same path that they had observed during the familiarization (inefficient approach). If we are right in our interpretation of the results of Experiment 3, according to which infants expect social group members both to faithfully copy the means actions of one another on the one hand, but also expect agents to realize their goals efficiently according to teleological stance, then this time they should be looking longer to inefficient actions in the test trials, given that the test agent is an outgroup member.

4.6.1 Participants

Participants were 24 infants (8 females). Mean age of the participants was 334.71 days (11 months) (range 321-350 days; 10;16 to 11;15, *SD* = 8.13 days). An additional 19 infants were tested but excluded from the final sample due to fussiness (*n*= 4), looking-times of less than 2 seconds per trial (*n* = 2), premature termination of a trial resulting from a lack of eye-tracking data for 2 consecutive seconds rather than looking away from the screen (*n* = 11). One baby removed the sticker on his forehead during the testing, and the software crashed for another baby during the testing.

4.6.2 Procedure and stimuli

The procedure was the same as in Experiment 3 except that a) the test agent was the outgroup agent, b) the spatial arrangement of the scene was slightly different as the outgroup test agent was positioned at the bottom while three ingroup agents were higher on the left side of the scene. The distance between the outgroup agent and the ingroup agent that was closest to it was the same as it was in Experiment 3. See Figure 4.12 for illustration of the stimuli.

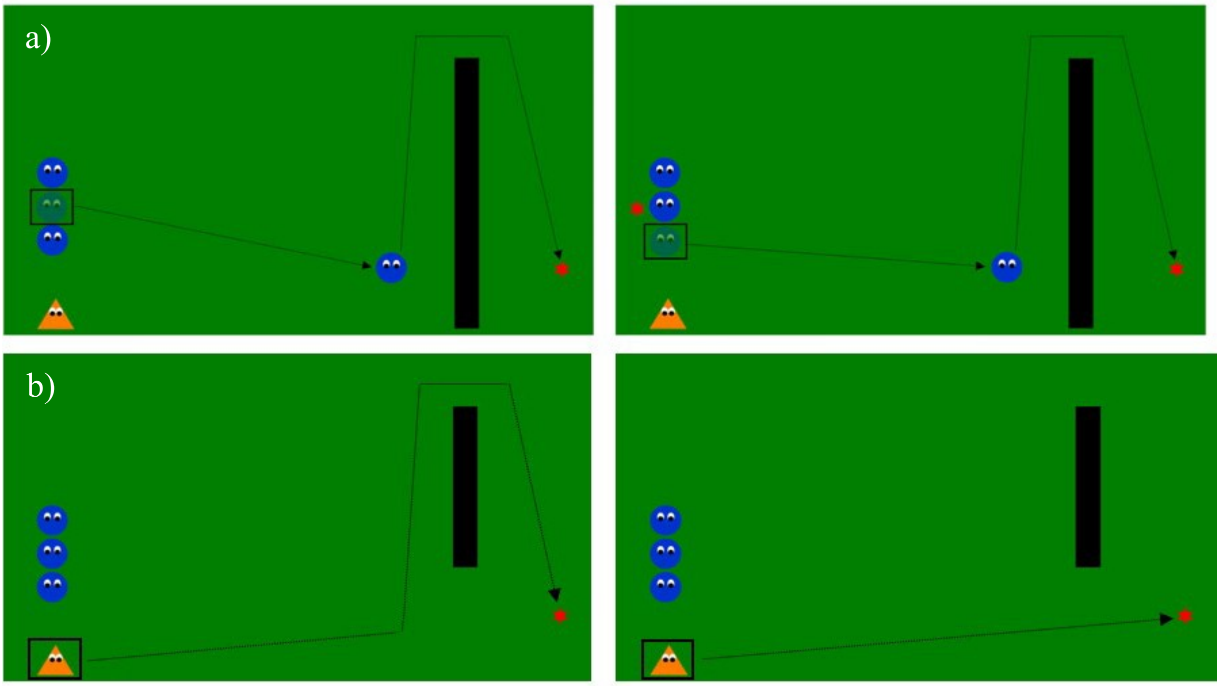

Figure 4.12. Schematic depiction of the stimuli showing: a) efficient actions performed by the two ingroup agents during the action familiarization phase, b) inefficient and efficient test events by the outgroup agent, respectively.

4.6.3 Results

4.6.3.1 Looking time analyses

Looking time data were log-10 transformed exactly as in the previous experiments. Data were collapsed across participant gender as preliminary analyses showed no significant interaction of this factor with test trials (F (1,22) = .18, p = .67, partial η^2 =. 008). There were 12 infants who viewed blue circles as the ingroup agents, while the remaining participants observed orange triangles as the ingroup agents. Which color and shape characterized the ingroup agents observed did not significantly interact with test trials (F (1, 22) = 2.2, p = .15, partial η^2 = .09). Half of the infants saw the ingroup agents moving first, while the other half saw the outgroup agent moving first during the group induction phase. Which group of agents moved first in the group induction trials did not interact with test trials F (1, 22) = 2.26, p = .15, partial η^2 = .09).

There were 12 infants who saw the efficient test action first, and 12 infants saw the inefficient test action first. Which type of test action was seen first significantly interacted with the test trials (F (1, 22) = 8.56, p = .008, partial η^2 = .28). Infants viewing the inefficient test trial in the first test pair they were shown, looked longer to inefficient test trials in comparison to efficient test trials (t (11) = -3.83, p = .003; Cohen's d = 1.12; Wilcoxon Signed Ranks test Z = -2.67, p = .005). Infants viewing the efficient test trial in the first test pair they were shown, did not look longer to inefficient test trials in comparison to efficient test trials. t (11) = -.31, p = .76, Cohen's d = .08; Wilcoxon Signed Ranks test Z = -1.09, p = .3). There was also a main effect of test trial (F (1, 22) = 10.81, p = .003, partial η^2 = .33). Infants overall looked longer to inefficient test trials (M = 35.37 s, SD = 27.52 s) than to efficient trials (M = 19.67 s, SD = 20.27 s). This difference was also statistically significant with a non-parametric test (a Wilcoxon signed-rank test on raw looking times, Z = -2.83, p = .004).

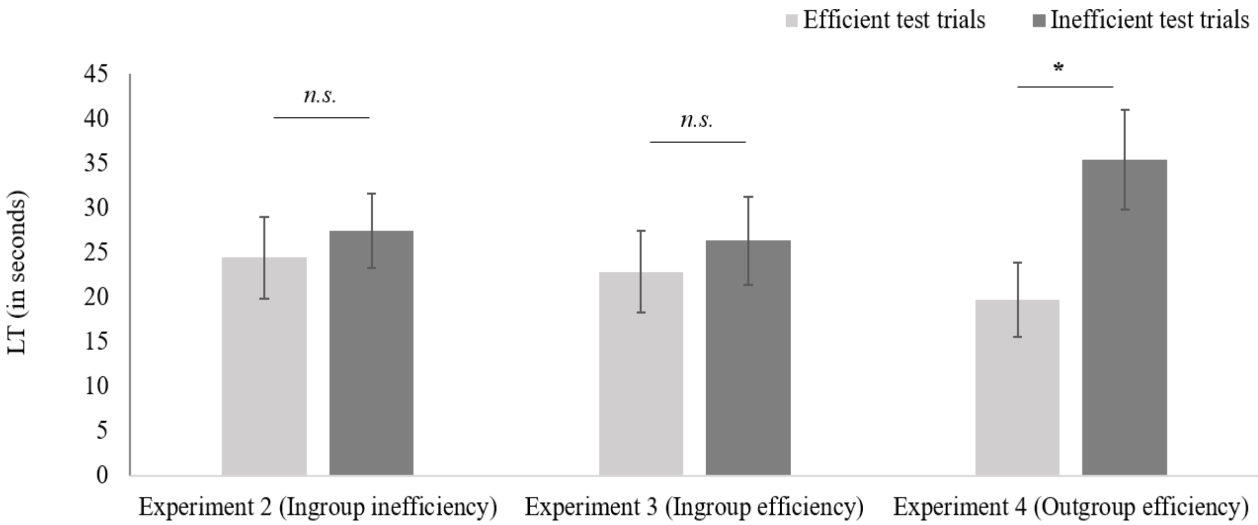

Figure 4.13. Mean looking times in seconds at the efficient and the inefficient test trials for infants in Experiments 2-4. Error bars show ±1 *SE*. Asteriks marks a statistically significant difference while *n.s.* marks no significant difference for the log-10-transformed data.

4.6.3.2 Pupil size analyses

For each participant the values within the time window chosen in the previous experiments (11700 ms – 17000 ms) were averaged and the individual averages were compared using both a repeated-measures t-test and its non-parametric equivalent. As depicted in Figure 4.14 there was a differential pupillary response induced by the two types of test events (t (21) = -2.37, p = .03, Cohen's d = .51; Wilcoxon Signed Ranks Test Z = -2.61, p = .007). Note that we had to exclude two infants due to missing data.

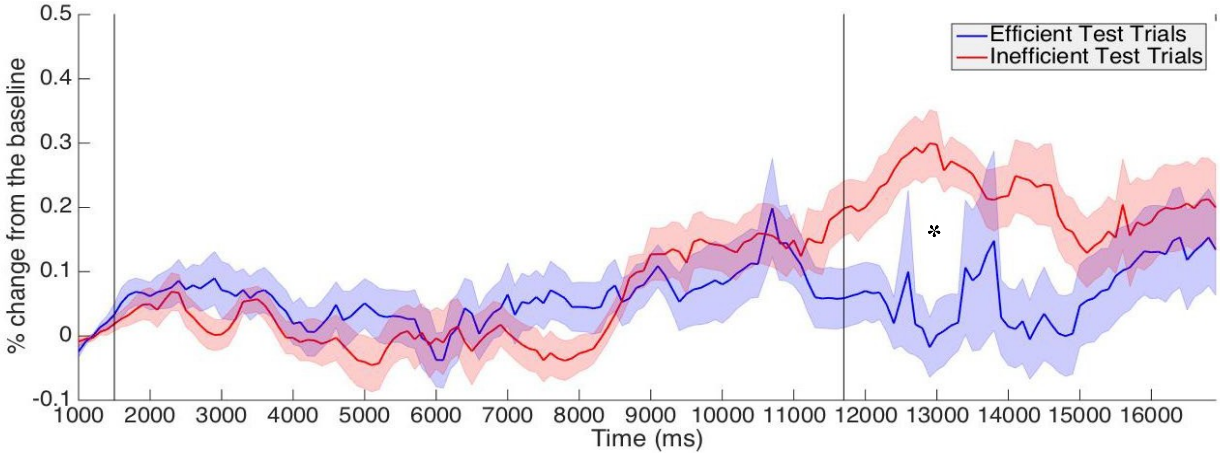

Figure 4.14. Temporal course of percentage of change in pupil size from the baseline for efficient and inefficient test trials in Experiment 4. The asterisk marks a statistically significant difference between the time points of 11700 ms -17000 ms.

4.6.4 Discussion

Having observed efficient approach actions by two ingroup members, which involved the group agents moving above the top of a long barrier in order to access the goal object, infants expected an outgroup agent to obtain the same goal efficiently when the barrier was shortened allowing the agent to access the object on a straight trajectory. Both looking time and pupil size measures reflected the increase in infants' attention to inefficient test trials in contrast to efficient test trials. This finding is in stark contrast to the findings of Experiment 3, where infants were presented with a novel ingroup agent, who had not been observed to act during the familiarization phase but was seen either acting efficiently or inefficiently in alternate test trials. It therefore appears that when the group membership information was not relevant for infants' predictions about the novel agent's future action, they did not have any trouble forming an expectation that was line with teleological stance. These findings suggest that social group information is what makes infants to attend to shared movement repertoire between agents, but as soon as this information was irrelevant, infants expected an efficient goal approach from a novel agent.

4.7 General discussion

Using a violation-of-expectation looking time paradigm we explored how infants form inferences about goal-directed actions of social agents belonging to groups. Specifically, we investigated how infants' expectations that agents pursue their goals efficiently, and that agents who belong to the same social group should behave alike interact. We predicted infants to be flexible in their interpretation of goal directed actions based on the familiarization events they viewed. We hypothesized infants to take into account inefficient approaches two different ingroup members performed repeatedly in the pursuit of a goal while predicting how the third ingroup member should act like in a similar scenario (Experiment 2). We expected infants to look longer when observing the third ingroup member perform a perceptually dissimilar action (even if it is an efficient goal approach) in contrast to when the third ingroup member carries out the same pattern of movement that its group members had been observed to perform (inefficient approach). Yet the findings showed that infants did not expect the third ingroup member to move alike as the other agents in its group did. This pattern presents a contrast with the findings of Powell and Spelke (2013). Critically, in our experiments we did not present the infants with the arbitrary movements that were performed by group members in the familiarization as it was in the study by Powell and Spelke (2013), but rather we presented infants with transitive events that involved an inefficient approach to a target object. We used inefficient actions based on the results of previous studies which documented that when the actions do not causally relate to the goal in the most efficient manner possible, older infants can encode such inefficient means chosen to attain a goal as a part of the goal hierarchy, treating them as a sub-goal to achieve the final goal (Kiraly, 2009b; Kiraly et al., 2013). However, it is a possibility that preverbal infants interpret the behaviors of agents acting in groups simply as shared movement patterns, and not as means actions. If infants were encoding the inefficient approaches of ingroup agents as mean actions, indicating the socially shared group sub-goal, then in Experiment 2 (when action familiarization was inefficient), they should have looked longer upon viewing a novel, but efficient action performed by another ingroup agent. In a similar vein, if they were not encoding the shared movement repertoire between the group members, then in Experiment 3 (when the familiarization action was efficient), they should have looked longer upon viewing the inefficient action performed by another ingroup agent. However, given the difficulties interpreting null results, it is an avenue for future research to probe whether preverbal infants represent inefficient actions of group members as shared movements or as means actions.

Nonetheless, there is evidence that infants do not construe the inefficient actions that is to be shared by same-group members as documented by an unpublished study (Powell, Schachner, & Spelke, 2014). In a similar design to their original study (Powell and Spelke, 2013), Powell and her colleagues presented goal-directed actions to 7.5- to 13.5-month-old infants. Infants viewed scenes in which there was a platform, along with two black boxes that turned purple in color upon contact. Two of the members from one group upon arriving on the platform jumped up and contacted the box that was positioned above the platform. Two other members from the other group slid on the platform leading them to contact the other box that was positioned on the right end of the platform. Similarly to our Experiment 2, in these scenarios there was no obstacle on the scene that would have required the agents to first go on the platform and perform respective jumping or sliding movements before contacting the boxes, as they could have just directly landed on the boxes to *operate* them. During the test phase, the last member of each group performed either the same action that matched the one performed by their group (consistent test) or they performed the action of the other group that did not match the action that their own group members had performed (inconsistent test). Infants' looking times to consistent and inconsistent actions did not differ from each other showing a similar pattern that we found in our Experiment 2.

Given the evidence suggesting infants' failure to represent inefficient approach actions as socially shared mean actions, it is a possibility that as soon as there is a potential external object that preverbal infants can construe as a goal that the agent pursues, young infants are inclined to apply their teleological inferences in interpreting and predicting the behaviors of others. Yet in Experiment 2 we presented infants with evidence that two agents from the same group do not pursue their goals efficiently. This alone might have disrupted infants' expectation of efficiency when predicting the action of the third group member. On the other hand, when infants were presented with evidence that two agents from the same group pursue their goals efficiently, they still did not look longer at test trials showing a third group member acting inefficiently (in contrast to test trials when the a third group member acted efficiently) (Experiment 3). These two findings suggest that infants do not only rely on the principle of rationality while forming expectations about how new agents should act, but also take into consideration the shared movement repertoires of agents that belong to the same groups. The findings of Experiment 4 strengthen this interpretation. Infants' looking time patterns revealed efficiency expectation only when the test agent was an outgroup member (Experiment 4).

We argue that this is so because infants' expectation for shared movement repertoire between the same-group members on the one hand, and their expectation that agents should act efficiently on the other, contradict each other. If shared movement repertoires between social group members had no influence on infants' expectation for efficient action, then we should have found no difference between Experiments 3 and 4. Yet looking time measurements revealed that this was not the case. When infants saw two ingroup members in the familiarizations going above the long barrier to obtain the goal object that rested on the other side (as in Experiment 3), they might have expected the other agent belonging to the same group as the familiarization agents to move the same way (i.e. going above the short barrier) but also to act efficiently (i.e. going under the short barrier to pursue a direct route towards the goal). These conflicting expectations might have been the reason for the null findings in infants' looking times. On the other hand, there was no reason for infants to generate contradictory expectations in Experiment 4 where they were tested for their predictions about the behavior of an outgroup agent whom they did not expect to move the same way as the agents belonging to a different group.

Yet we should note that the pupillary measures we have collected in Experiment 3 revealed infants' enhanced attention in response to inefficient action trials just like it was the case in Experiments 1a, 1b and 4. This was so even though infants did not differentiate efficient and inefficient tests actions in their cumulative looking times in Experiment 3. Considering that cumulative looking time and pupillary measures were collected in different time points, it is a possibility that infants' initial response to inefficient actions as revealed by their increased pupil size during the test reflected a local and momentary reaction that did not carry over to the whole time period for which we have collected the cumulative looking time measure. Pupil response can be a proxy for their initial teleological expectation, but within the time frame of the test trials this expectation could be counteracted by accumulating evidence supporting their alternative expectation of exhibiting a shared movement repertoire that characterize same-group members.

Furthermore, our post-hoc pupillary analysis for the familiarization events in Experiments 2 and 3 could be interpreted as constituting evidence for infants' goal attribution as a function of action efficiency. Given the inefficient approaches in the familiarization scenes of Experiment 2, it is likely that infants did not generate any goal attribution during this phase and there was consequently no expectation they could had formed to be violated in the test. In contrast, infants in Experiment 3 were able to attach a goal to the efficient action they viewed in the familiarization, and their differential pupil response to inefficient trials during the test reflected the violation of this goal attribution, acting as a more sensitive measure than the cumulative looking time. However, since we designed this series of experiments as violation of expectation looking time studies, it is an avenue for future research to follow up the nature of the findings we obtained through our exploratory pupillary measure.

We should also call the reader's attention to the fact that unlike in the paradigm of Powell and Spelke (2013), we presented social groups to infants with three animated characters as ingroup agents and another animated character as an outgroup agent. The original study instead made use of six characters making up two groups. Since in Experiments 2-4, we always presented infants with the actions that the majority of the ingroup members performed, and later tested how they expected a novel agent, who either belonged to the majority group or not, would approach the goal object, one can find it likely that the social group information we provided the infants with was not as contrastive as it was in Powell and Spelke (2013). However, Powell, Schachner and Spelke (2014) using the same number of characters making up two groups as in Powell and Spelke (2013), also reported that infants do not distinguish different sub-efficient means that two groups of agents performed in their goal retrievals.

Furthermore, Powell and Spelke (2013) documented that when all six agents were distinctive in their physical appearance from each other but still danced in synchrony in groups of three, 8-month-old infants were able to recognize group-consistent and group-inconsistent behaviors of animated characters (i.e. landing on the same object as their group members or on a different object). Hence it is a possibility that the synchronous dancing events infants in our paradigm were also presented with is the reason why infants did not show robust efficiency expectation when the test agent was an ingroup member. It might be the case that the synchronous dancing event by itself enabled infants to expect the group of characters, who earlier moved together in synchrony with each other, to share the same movement repertoires later. It is not clear whether infants would still show sensitivity to the shared movement repertoires of agents if the dancing event were removed from the group induction scenes, and if the group membership were indicated by only shared visual features. It is an avenue for future research to explore how different cues of social group membership enable infants to represent agents to belong to social groups and to expect similar behaviors to be performed by them in changing situations.

In sum our looking time findings show that pre-verbal infants guide their inferences about the agent's behavior based on two expectations: teleology and shared movement repertoire of group members. Our findings in Experiments 2, 3 and 4 suggest that at the age of 11 months these two expectations might contradict each other. In fact, Spelke (2016) argued that infants' representations of agency and social beings reflect two encapsulated systems, and these remain distinct in young infants. Infants interpret others' actions either as instrumental, guided by external goals or as social, guided by the social forces promoting conformity, but not both at the same time (Spelke, 2016). Our findings showing the effect of shared movement repertoire on 11-month-old infants' efficiency expectations in goal directed actions suggest that infants' representations of agents and social beings are not guided by two independent systems, yet it also appears they are still not fully integrated either in preverbal infants. Spelke (2016) proposed that only when infants start to comprehend or produce linguistic expressions that refer to social goals (e.g. "Look, Simon, an apple") can these two systems become combined creating a new system of knowledge to represent agents who act instrumentally in the physical world and also who interact with other social beings in causally opaque ways.

This hypothesis for the necessity of the combinatorial capacity provided by language to allow infants to represent actions both as instrumental and social at the same time has not been empirically tested yet (but see Schmidt, Rakoczy & Tomasello, 2019) and it is a topic beyond the scope of this chapter. However, whether older infants could integrate teleological inferential principles with their expectation of social behavior that is shared among the agents belonging to the same social group in guiding their predictions about the behaviors of others is an interesting question for future research. Would older infants expect an agent to behave in similar inefficient ways as her group members did or would they expect her to pursue her goal efficiently independently of what her group members acted like before? Similarly, would 11-month-olds upon being familiarized with inefficient action approaches of two ingroup members expect an outgroup agent to pursue the goal object efficiently? Or if one is to remove the goal object from the scenes altogether, would 11-month-old infants then expect group members to act alike?

If we are right in our interpretation that infants expect agents in groups to imitate movement paths, and this conflicts with their expectation of efficient action (hence null results with Experiments 2 and 3) then blocking the input conditions that could activate the teleological action interpretation (removing the goal from the scene altogether) or blocking the input conditions that could active shared group movement repertoire (testing an outgroup agent or removing group cues altogether) would continue yielding different pattern of findings among preverbal infants. So far, the evidence we have points to the likelihood that preverbal infants expect agents both to move the same way as their group members but also to pursue their goals efficiently, while the latter appears a more robust expectation. It is an important question for future research when and how these two systems become fully integrated, allowing infants to form expectations in line with teleological interpretation of the actions and also expect shared movement repertoires of social groups, yielding an understanding that actions could be motivated by the conventions of social groups over and beyond the instrumental effects they seem to achieve. After all, representing others' actions not only as constrained by the physical environment, but also by the normative expectations socially stipulated and shared by the social group they are in might be foundational to understanding cognitively opaque cultural practices of groups.

Chapter V. Acquiring sub-efficient and efficient variants of novel means by integrating information from multiple social models in preschoolers

5.1 Introduction

Agents tend to pursue their goals through efficient means, but in certain social contexts they often preserve customary, conventionalized and normative means actions of their cultural groups, which tend to deviate from efficiency (Gergely & Csibra, 2006; Legare & Nielsen, 2015). Such sub-efficient action routines are socially transmitted and form part of the repertoire of cultural knowledge shared by members of a child's community. Acquiring such sub-efficient yet culturally shared routines is, therefore, an important developmental task that significantly contributes to children's socialization to become competent members of their cultural group. Indeed, preschool-aged children readily learn novel sub-efficient means from social models as documented by so-called "over-imitation" [5] studies (e.g. Horner & Whiten, 2005; Lyons, Damrosch, Lin, Macris, & Keil, 2011; Lyons, Young, & Keil, 2007; for a review Hoehl, Keupp, Schleihauf, McGuigan, Buttelmann, & Whiten, 2019). It is less clear however, what enables learners to preserve such actions in their repertoire despite the high likelihood of encountering more efficient alternative means of achieving the same goal.

Exposure to efficient alternatives can happen through contact with agents, who either do not possess the shared knowledge of the child's cultural group (Olah, Elekes, Brody, &

[5] We find the term "over-imitation" misleading as it suggests that children imitate what they should not imitate. However, we would argue that children's tendency to "over-imitate" cognitively opaque actions demonstrated by knowledgeable adults is, in fact, a "smart" cultural learning strategy based on the assumption that such sub-efficient means *should* be imitated precisely because they are evaluated as culturally relevant (despite their opacity to the learner).

Kiraly, 2014), or who do so, but on particular occasions pursue the goal more efficiently (Burdett, McGuigan, Harrison, & Whiten, 2018; Evans, Laland, Carpenter, & Kendal, 2018; Hoehl, Zettersten, Schleihauf, Gratz, & Pauen 2014). After all, the sub-efficient culture-specific means action often co-exists in a cultural community with more efficient variants that are also considered acceptable in every-day situations. For example, when you are invited to a formal dinner you may be expected to wear a tuxedo and follow rules of etiquette prescribing that you *ought to* consume your dinner in a culturally standardized *proper* manner performing sub-efficient, often opaque and arbitrary, normative action routines. However, when you are having lunch on the go it may be quite acceptable to adopt a more relaxed everyday style of eating that is not constrained by conventionalized cultural prescriptions, allowing you to achieve your goal through more efficient action alternatives.

Given the frequent co-existence of sub-efficient forms of actions along with more efficient alternative means, how do naïve learners integrate such conventional action routines and their efficient alternatives? This is not a trivial learning problem for a young cultural learner to solve. It has been argued, that cultural means actions, i.e. ways of doing things in various cultural groups, are often inherently *cognitively opaque* to the naive observer, such as a child learner (Csibra & Gergely, 2011; Gergely & Csibra, 2006; Hernik & Gergely, 2015; Hoehl et al., 2019). In other words, e.g. the ultimate goal of the action sequence or the factors assumed in the cultural group to constrain and influence the action steps, may not be readily apparent to the child. Thus, even though already young infants expect goal-directed actions to be efficient (Csibra, Gergely, Biro, Koos, & Brockbank, 1999; Gergely & Csibra, 2003; Gergely, Nadasdy, Csibra, & Biro, 1995), the primary mechanisms supporting acquisition of the means actions customary in the child's cultural group should not rely on efficiency considerations, since the teleological and casual structure behind such actions often will not be accessible to the child. Indeed, children tend to imitate the casually irrelevant action steps, even if they know them to be unnecessary or if they are unsure of their causal role (Kenward, Karlsson, & Persson, 2011; Keupp, Behne, & Rakoczy, 2013). So, what kind of information can young children rely on to acquire cognitively opaque cultural means and to either maintain them or not, after learning about more efficient and acceptable action alternatives?

Several findings suggest that certain social signals may help children adopt the efficient alternative to the demonstrated sub-efficient action routine, indicating how flexible children's imitation in fact is (review in Over & Carpenter, 2012a; 2012b). For example, if demonstration was accompanied by outcome-oriented language cues (e.g. "I am going to make a necklace"), as opposed to convention-oriented language cues (e.g. "Everyone always does it this way"), preschoolers were more likely to prioritize efficient attainment of the goal despite the demonstrated sub-efficient means (Clegg & Legare, 2016; Hermann, Legare, Harris, & Whitehouse, 2013; Legare, Wen, Hermann, & Whitehouse, 2015). Moreover, when the demonstration of a sub-efficient action routine was verbally framed so as to emphasize the particular way in which the outcome should be attained (i.e. "daxing" as opposed to "ringing the bells") children were more likely to protest if a third party (i.e. a puppet) omitted the sub-efficient mean action (Keupp, Behne, & Rakoczy, 2013; Keupp, Behne, Zachow, Kasbohm, & Rakoczy, 2015). Altogether these studies document that children can flexibly engage in conventional reading of the task and reproduce the process faithfully (and also expect third parties to reproduce the process faithfully), or in instrumental reading of the same task and prioritize the goal-outcome in their imitation, based on the different social contextual cues.

Critically, even though children can acquire new cultural knowledge from 3rd-person observations of others' actions (Phillips, Seston & Kelemen, 2012), cultural knowledge transmission often takes place in 2nd-person ostensive communicative context, where an adult social partner uses ostensive signals to address and demonstrate to the child the new and relevant cultural skill or knowledge to be acquired (Csibra & Gergely, 2006; 2009; 2011;

Egyed, Kiraly, & Gergely 2013; Gergely & Jacob, 2012; Kiraly, Csibra, & Gergely, 2013). It has been shown that young children are sensitive to ostensive-communicative demonstration contexts, which lead them to establish stable cultural representations of novel means-actions. Furthermore, the acquired functions of novel tools ostensively demonstrated, appear to be relatively resistant to change through counter-examples and tend to "survive" competition coming from later evidence (Butler & Markman, 2012; Hernik & Csibra, 2015).

Thus, whether a sub-efficient action routine already acquired is going to be maintained or rather modified as a result of encountering more efficient action alternatives may depend on whether the efficient alternatives were communicatively or non-communicatively presented to the child. Hoehl and her colleagues (2014) addressed this question by investigating the influence of ostensive communicative signaling on cultural learning in 5-years-old preschoolers. Consistent with the burgeoning literature (Horner & Whiten, 2005; Lyons et al., 2007; 2011; Nielsen & Blank, 2011), 5-year-olds faithfully imitated the causally irrelevant action steps that the model performed on the box before retrieving the token inside with the last causally efficacious action. Yet, when this sub-efficient routine was followed by a second model's demonstration of a more efficient alternative (the last action step only), children flexibly adopted this efficient action. However, they did so only in the condition, where the efficient alternative was presented within an ostensive communicative context. In another condition where no ostensive signals were provided by a non-communicative model, who presented the more efficient alternative, children did not modify their already acquired sub-efficient action routine.

However, apart from the communicative versus non-communicative source of novel information the cultural learning process may also be influenced and guided by another relevant informative factor that is the focus of our current study: children may preferentially rely on information coming from agents, who they can identify as expert sources of the cultural knowledge repertoire shared in the child's cultural group (Olah et al., 2014). One informative signal that the child can use to identify reliable sources of cultural knowledge is shared language. Even very young infants prefer speakers of their native language (Kinzler, Dupoux, & Spelke, 2007; Shutts, Kinzler, McKee, & Spelke, 2009). A recent EEG study indicated that such preferences are driven by infants' expectation that they are likely to receive relevant information from agents speaking their own language (Begus, Gliga, & Southgate, 2016). Yet, direct empirical evidence on young children's propensity to selectively evaluate and learn from the speakers of their own language is rather limited (Buttelmann, Zmyj, Daum, & Carpenter, 2013; but also see Howard, Henderson, Carrazza, & Woodward, 2015). One possible reason is that young children may interpret shared language not as a categorical cue, that renders them either willing or unwilling to learn from a model dependent on whether she speaks child's own language or not, but rather as a nuanced signal, which informs the likelihood that the model possesses the knowledge of the child's cultural group.

In this chapter we test the hypothesis that in preschool children maintaining the already acquired sub-efficient means action is facilitated by a selective evaluation of the model, presenting the more efficient alternative. Based on recognizing that the model is speaking the shared language (the one used in the child's cultural group), the child may be able to infer that the model is likely to possess the knowledge repertoire shared in the child's own cultural group. If so, this will influence whether the child adopts and integrates the alternative efficient means or disregards them as culturally non-relevant.

5.2 Experiment 1

In experiment 1 we employed a between-subject design, in which children received two ostensive demonstrations performed by two different models, each showing the child a different way of operating a novel apparatus to retrieve a sticker from it. Our procedure closely resembled Hoehl et al.'s (2014) study and we targeted the same age group in our sample. However, the models in Hoehl's et al.'s (2014) experiment displayed verbal ostensive signals before demonstrating their novel means actions to the child. In contrast, in our procedure neither model addressed the child verbally before or during their respective demonstrations. Furthermore, both models used only *non-verbal* ostensive and referential gestures before demonstrating either the sub-efficient action routine or the more efficient action alternative to retrieve the sticker from the same apparatus (such as establishing eye-contact and smiling at the child, referentially pointing to the apparatus, and looking back-and-forth between the child and the apparatus). This way both action demonstrations were equally communicative and involved ostensively addressing the child in a second-person manner before the demonstrations of the novel action sequence. On the other hand, information about whether the model was speaking in their own language or in a foreign language unfamiliar to them was manifested only in an implicit manner in the "over-heard" phone conversation that the slowly approaching model was engaged in and that was clearly not addressed to the child (in fact, it was accessible to him only through third-person observation). This initial period of implicit observational access to the language cue ended prior to the demonstrations when the model put her phone down and turned for the first time to the child addressing her with second-person ostensive communicative (though non-verbal) signals to introduce the demonstration phase.

The first model always demonstrated a *sub-efficient* action routine in retrieving a sticker (by performing a series of superfluous actions on the apparatus before presenting the only causally relevant action that was necessary to obtain the sticker), and the second model always presented a more *efficient* action alternative (by performing only the causally relevant action). Children were given the apparatus after each demonstration to operate it themselves. Before these action demonstrations children witnessed the model speak either the child's own language or a foreign language.

In the *same* + *foreign* condition, the first model entered the room speaking on her cellphone to someone in the child's own language. Similarly, when the second model entered, she was also speaking on her cellphone to someone, but in a foreign language. We expected children in this condition to reproduce the novel sub-efficient action routine ostensively demonstrated to them by the first model. We also expected that children will continue to reproduce the initially demonstrated sub-efficient action sequence even after having observed the second model's presentation of the more efficient alternative means action.

In order to rule out the possibility that this predicted pattern of imitation in the *same* + *foreign* condition was merely due to children's reluctance to learn from a foreign-language speaker (Buttelmann et al., 2013; Howard et al., 2015), we also tested another group of children in a *foreign* + *same* condition, where the first model spoke the foreign language on the phone and the second model spoke child's own language. Here we predicted that even if during the first imitation period children reproduce the sub-efficient action routine of the first model, during the second imitation period they would later adopt the more efficient alternative of the second model.

5.2.1 Participants

There were 16 5-year-olds in the *foreign* + *same* condition (5 females) (M_{age} = 65.43 months, *SD* = 3.86 months, range: 60.68-71.88) and 16 5-year-olds in the *same* + *foreign* condition (10 females) (M_{age} = 66.03 months, *SD* = 4.14 months, range: 60.71-71.81). There were 16 6-year-olds tested in the *foreign* + *same* condition (6 females) (M_{age} = 76.82 months, *SD* = 2.78 months, range: 73.13-81.83) and a further 16 6-year-olds participated in the *same* + *foreign* condition (10 females) (M_{age} = 78.24 months, *SD* = 3.39 months, range: 72.87-84.82)

A further 11 5-year-olds were tested but excluded from all analyses due to not acting on the sticker dispenser in the first imitation round (n = 1) or in either imitation rounds (n = 2), not watching the demonstration (n = 2), seeing the resetting of the sticker dispenser (n = 1), and for experimenter error (n = 5). An additional two 6-year-olds were also tested but excluded from all analyses due to seeing the resetting of the sticker dispenser. All participants in the final sample were monolingual (as established based on parental reports during consent), speaking only Hungarian as their native language, and were recruited from local kindergartens in Budapest inner city area. Parents were contacted by the head teacher in the participating kindergartens to give written consent. The study was approved by the United Ethical Review Committee for Research in Psychology, Hungary, and conducted in accordance with the Declaration of Helsinki.

5.2.2 Materials

The mini-bowling game used for the warm-up phase involved two cups placed on top of each other with two stickers inside, and a tennis ball. There were three identical 10 cm x 10 cm sticker sheets for the stickers collected, one for the child and one for each model. The sticker dispenser (Fig. 1) was a transparent plexi-glass box with a wooden platform inside, which could be pulled out by using a tool provided. On the platform there was a small container with four stickers. Another small plexi-box with an opening on its top and a button inside was attached to the right side of the container. There was also a wooden lever glued on the top of the container box.

5.2.3 Procedure

All participants were tested in their own kindergartens in a quiet room. A female research assistant acted as the experimenter and escorted each participant to the testing room.

Warm-up phase. The experimenter introduced the scene to the participant by saying, “Oh there are two games here, both for getting stickers: let’s play with this one first” and directed the child’s attention to the mini-bowling game, which consisted of rolling a tennis ball to knock down the cups with stickers in them. The game was played to ensure that the child becomes comfortable with the testing room. Having retrieved the first sticker when playing the game, the experimenter handed over a sticker-sheet to the child and showed how to place the sticker on the sheet. She announced: “Try getting the other sticker on your own now, I am a bit busy. I need to finish reading these reports. Let me know when you get the other sticker and when you are done placing it on your sheet”. She then sat down on a chair in the corner of the testing room and pretended to read. This was done to ensure that the children understood that the experimenter was busy and did not pay attention to what they were doing.

After this warm-up game, the experimenter said that she was going to keep the child’s sticker-sheet and that other people would be coming to the room to also get a sticker for themselves by using the sticker-box. The experimenter instructed the child to sit on a mat next to the box and to watch closely. She then went back to reading while sitting in the same corner of the testing room close to the door, facing away from the child and the apparatus.

The first demonstration phase. The first model came into the room speaking to someone on her cellphone for about 15 seconds on a topic unrelated to the procedure (see Appendix 3). As she was sitting down in front of the sticker dispenser, the model hung up the phone, looked and smiled at the child, showed her own sticker-sheet and pointed at the stickers in the dispenser to indicate her intention to get a sticker. There were three irrelevant actions that the model was performing on the apparatus (manipulating the lever [a], pushing the button inside the small box with the tool [b], tapping the tool on the side of the apparatus [c]), and one relevant action (pulling the platform out with the tool [d]) that was necessary to obtain the sticker. Sub-efficient action demonstration consisted of all four actions, in the order depicted in Figure 5.1 [a-d].

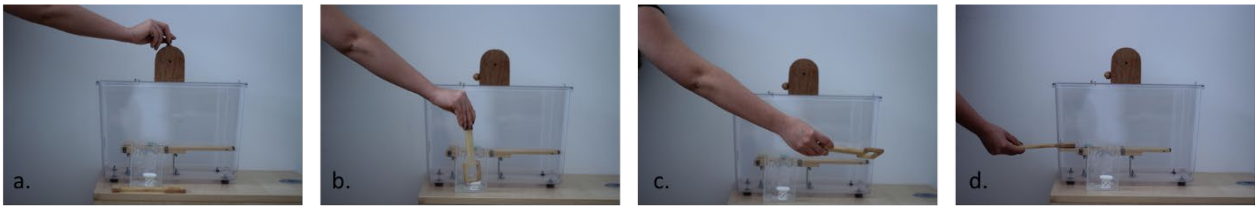

Figure 5.1. Sticker dispenser and three irrelevant actions, and one relevant action in attaining the sticker.

Once the model successfully retrieved the sticker from the dispenser, she again looked and smiled at the child, put her sticker-sheet on the floor and placed the retrieved sticker on her sheet. This conspicuously signaled the end of the demonstration by the experimenter.

The experimenter then approached the child and oriented her or him towards herself and away from the apparatus and the model while saying "Oh how old are you? Are you [child's age (e.g. five)]? Let's count together to [child's age]", thus allowing the model to reset the sticker dispenser without the child watching. As soon as the model left the room, the experimenter pointed at the dispenser and said, "It is your turn now". Then she went back to the corner and sat down on her chair, facing towards the door and away from the child. This was the beginning of the first imitation round. If the child did not act on the sticker dispenser for about 30 seconds, the experimenter turned around and asked, "Do you want to get a sticker?" as a further prompt for the child to act on the dispenser. If the child did not act on the dispenser for approximately 90 seconds, the imitation round was terminated.

If the child successfully retrieved a sticker in the first imitation round, the experimenter returned the sticker-sheet while orienting the child away from the apparatus, so that the experimenter could reset the dispenser without being seen by the participant. She then said that there were more people coming to get a sticker from the box and that the child should watch closely, and she again sat back in her chair in the corner pretending to read.

The second demonstration phase. The second model came in, again talking on her mobile phone (see Appendix 3). Everything was the same as it was in the first demonstration, except that this time the model attained the sticker efficiently. The efficient action demonstration consisted only of pulling the platform out with a tool (Figure 5.1.d).

After the model left the room, the child was again prompted to get a sticker by the experimenter with the instruction "It is your turn now". This was the beginning of the second imitation round.

If the child was unable to retrieve a sticker during one or both imitation rounds despite acting on the sticker dispenser, she or he was excluded. This exclusion criterion ensured that all children were equally successful in obtaining the sticker in both imitation rounds. At the end of the testing session each child received four stickers in total (two from the mini-bowling game, two from the sticker dispenser) regardless of their performance. Two models were white female post-graduate students of two different nationalities with similar physical features (white skin tone, dark hair, brown eyes) and they were native speakers of the respective languages. All the testing sessions were videotaped for later coding.

5.2.4 Coding

The number of irrelevant actions the children imitated was coded for each imitation round (range: 0-3) as an imitation score. Approximately 30% of the data (10 children) were coded by a second coder naïve to the condition and the hypotheses of the study. Inter-coder reliability was high for 5-year-olds (Cohen's Kappa = .94, $p < .001$) and perfect for 6-year-olds (Cohen's Kappa = 1, $p < .001$). See Appendix 3 for the coding of causally relevant action.

For all participants, we also coded the amount of time they looked at the demonstration to check whether the familiarity or novelty of the model's language might have affected children's overt attention to the demonstrations.

5.2.5 Results

Data were collapsed across participant gender as preliminary analyses showed no significant main effects nor interactions of gender with condition and imitation round.

A 2x2x2 (condition x imitation round x age) repeated measures ANOVA on children's imitation score revealed no statistically significant three-way interaction but a main effect of imitation round, F (1, 60) = 94, p < .001, partial η^2 = .61. There was also a statistically significant interaction between imitation round and age, F (1, 60) = 3.78, p = .03, partial η^2 = .08. There was no main effect of condition for either imitation round (ps > .05) Please see Figure 5.2.

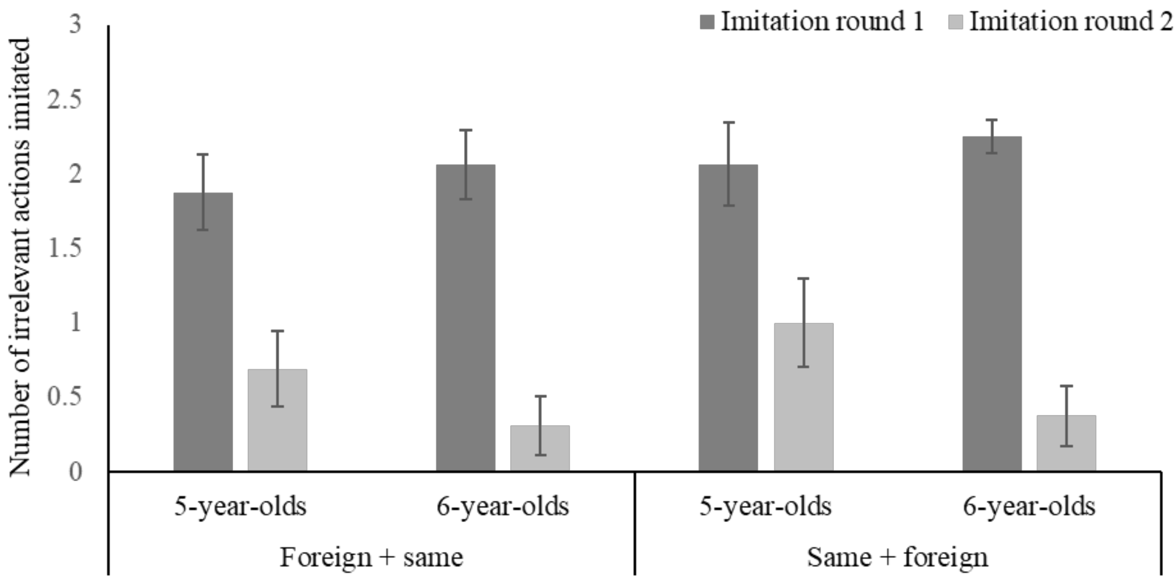

Figure 5.2. Mean number of irrelevant actions imitated by 5- and 6-year olds in each imitation round in two conditions. Error bars show ±1 *SE*.

Planned comparisons revealed that children's imitation scores decreased significantly from the first to the second imitation round both in *foreign + same* condition ($M_{foreign}$ = 1.97, SD =.97; M_{same} = .5, SD = .92), t (31) = 6.98, p < .001 Cohen's d = 1.23; Wilcoxon signed ranks test, Z = -4.17, p < .001, and in the *same + foreign* condition (M_{same} = 2.16, SD = .85;

$M_{foreign}$ = .69, *SD* = 1.06), *t* (31) = 6.42, *p* < .001, Cohen's *d* = 1.13; Wilcoxon signed ranks test, *Z* = -4.03, *p* < .001. There were no differences in 5- and 6-year old's imitation scores in the first round, *t* (62) = .41, *p* = .26, Cohen's *d* = .21; Mann-Whitney *U* = 495, *p* = .77. However, in the second imitation round imitation score of 5-year-olds was higher (*M* = .84, *SD* = 1.11) than 6-year-olds (*M* = .34, *SD* = .79). This difference was statistically significant with a parametric test, *t* (62) = 2.08, *p* = .042, Cohen's *d* = .52 but not with its non-parametric equivalent, Mann-Whitney *U* = 392.5, *p* = .05.

Demonstrations from the same- and the foreign-language speaker were attended to equally well (all medians = 100% of the demonstration duration for two age groups).

5.2.6 Discussion

Consistent with the literature (Hoehl et al., 2014; Schleihauf, Graetz, Pauen, & Hoehl, 2018), the 5- and 6-year-olds readily imitated the clearly sub-efficient action sequence that was ostensively demonstrated to them by the first model who showed them how to retrieve a sticker from the novel apparatus. Notably, children in both age groups tended to reproduce most steps of the sub-efficient action routine that was ostensively demonstrated to them by an unfamiliar adult regardless of whether she was a speaker of their own language or a foreign language. This indicates that the over-heard foreign language did not result in a categorical interpretation of the speaker as a model not to be imitated.

Critically, and contrary to our hypothesis, once the efficient alternative was demonstrated in the second round, children adopted the efficient means action *both* when it was presented by the speaker of their own language and when it was presented by a foreign-language speaker. Below we consider some possible reasons for this.

On the one hand, the age group (5-6-year-olds) we chose may have been too *old* for testing our hypothesis. Older preschoolers' experience with multilingual speakers in their social environment may have made it acceptable for them that an over-heard conversation with an unseen interlocutor in a foreign language does not rule out the possibility that the person is *also* a speaker of child's own language.

Moreover, by this age children might have already learned to put relatively less weight on others' language when deciding whether to acquire cultural information from them. Previous research indeed documented that incidental exposure to linguistic diversity was a factor influencing children's propensity to learn from a model speaking a foreign language (Howard, Carrazza, & Woodward, 2014). Also, as recent research shows, older preschoolers do not always choose to imitate the sub-efficient means when these were pitted against the efficient ones simultaneously (Burdett et al., 2018; Evans et al., 2018). Considering the trend in our data, which indicated that 6-year-olds more readily adopted the efficient alternative in their second imitation round in contrast to 5-year-olds, independently of the model's language, it is a possibility that older children might be more likely to employ the efficient alternative to complete the task.

Yet, the above results do not rule out that children – especially with less relevant social experience - can be sensitive to the language of a model as an informative cue to her cultural competence. To explore this possibility, in experiments 2 and 3 we tested 4-year-olds. Children at this age can already reliably imitate complex target behaviors (Kenward, Karlsson, & Persson, 2011) and can rely on contrastive information that can be diagnostic of social groups (Chalik, Rivera, & Rhodes, 2014).

5.3 Experiment 2

Since our procedure did not involve any direct verbal communication between the models and the participant during the demonstration phase (as it was the case in Hoehl et al.'s paradigm, 2014, with 5-year-olds) we first wanted to ensure, that in 4-year-olds the non-verbal ostensive demonstrations also elicit high imitation of sub-efficient actions followed by reduction upon encountering the non-verbal efficient alternative. Thus, experiment 2 had procedure exactly like experiment 1, except that both models were speakers of the child's own language (cf. *pedagogical-then-pedagogical* condition in Hoehl et al., 2014).

5.3.1 Participants

Participants were 16 4-year-olds (7 females) (M_{age} = 53.51 months, *SD* = 3 months, range: 49.18-57.7 months). A further 3 4-year-olds (2 females) were tested but excluded from all analyses due to not acting on the apparatus in the first imitation round. All participants were monolingual, speaking only Hungarian as their native language and were recruited from local kindergartens in Budapest inner city area.

5.3.2 Design and procedure

The procedure was the same as in experiment 1 except that both models were speakers of child's own language. Both models were white female post-graduate students of the same nationality with similar physical features (white skin tone, dark hair, brown eyes), and were locals of the same city speaking with the same accent.

We dubbed this condition as *same* + *same*. We counterbalanced which model attained the sticker sub-efficiently. Half of the children had seen one model acting sub-efficiently in the first demonstration phase, while the other half had seen that model acting efficiently in the second demonstration phase.

5.3.3 Coding

The number of irrelevant actions the children imitated was coded for each imitation round separately (range: 0-3). Data from five children were also coded by a second coder who was naïve to the condition and the hypotheses of the study. Inter-coder reliability was perfect (Cohen's Kappa = 1, $p < .001$).

5.3.4 Results

Data were collapsed across participant gender and model identity, as preliminary analyses showed no significant main effects nor interactions of these factors in the imitation round.

Mean number of irrelevant actions imitated was higher in the first (M = 1.94, SD = 1.29) than in the second (M = 1.13, SD = 1.36) imitation round, t (1, 15) = 2.66, p = .02, Cohen's d = .66, Wilcoxon signed-rank test Z = -2.23, p = .03 (Figure 5.3). Children attended equally well to both demonstrations (both medians = 100% of the demonstration).

5.3.5 Discussion

In experiment 2 we showed that in the first imitation round 4-year-olds tended to imitate sub-efficient means of achieving the goal demonstrated by a same-language speaker. Crucially, the model's demonstration was ostensive, but fully non-verbal as it was in experiment 1. Furthermore, in the second imitation round, after observing a demonstration of an efficient action alternative presented again communicatively yet non-verbally by another same-language speaking model, 4-year-olds adopted the efficient alternative and imitated the superfluous causally irrelevant actions of the previously acquired sub-efficient action routine significantly less. These results with 4-year-olds are in line with those reported by Hoehl et al. (2014) with older preschoolers and extend them to a context of non-verbal communicative demonstrations with younger children. After validating the experimental procedure with 4-year-olds, we proceeded to test our main hypothesis in this age group in experiment 3.

5.4 Experiment 3

Experiment 3 had the same structure as experiment 1, except that participants were a group of 4-year-olds In line with our earlier predictions, we expected 4-year-old children in the *same + foreign* condition to continue imitating the already acquired sub-efficient routines, even after observing an efficient alternative means action demonstrated by a foreign-language speaker. In contrast, we expected children in the *foreign + same* condition to adopt the efficient alternative presented by the speaker of their own language.

5.4.1 Participants

There were 16 4-year-olds in the *foreign + same* condition (8 females) (M_{age} = 54.31 months, *SD* = 4.02 months, range: 48.88-58.9) and 16 4-year-olds in *same + foreign* condition (6 females) (M_{age} = 54.5 months, *SD* = 3.7 months, range: 48.13-59.13).

A further 25 4-year-olds (15 females) were tested but excluded from all analyses because they did not act on the sticker dispenser at all in either one (n = 1) or both imitation rounds (n = 4), did not retrieve the sticker from the sticker dispenser in one (n = 10; $n_{same+foreign}$ = 4, $n_{foreign+same}$ = 6) or both imitation rounds (n = 4; $n_{same+foreign}$ = 2, $n_{foreign+same}$ = 2), did not watch the demonstration (n = 2), saw the resetting of the sticker dispenser (n = 1), or due to camera failure (n =1), or experimenter's error (n = 2).

All participants in the final sample were monolingual, speaking only Hungarian as their native language, and were recruited from local kindergartens in Budapest inner city area.

5.4.2 Design and procedure

The design and procedure were the same as in experiment 1. The models were the same models who performed the demonstrations in experiment 1.

5.4.3 Coding

The number of irrelevant actions the children imitated was coded for each imitation round separately (range: 0-3) as an imitation score. Approximately 30% of the data (10 children) were coded by a second coder who was naïve to the condition and the hypotheses of the study. Inter-coder reliability was perfect (Cohen's Kappa = 1, p < .001). If the camera-angle allowed ($n_{same+foreign}$=14, $n_{foreign+same}$=15), we also coded the amount of time participants looked at the demonstrations to check whether the familiarity or novelty of the model's language might have affected children's overt attention to the demonstration.

5.4.4 Results

Data were collapsed across participant gender, as preliminary analyses showed no significant main effects nor interactions of this factor with condition and imitation round.

A 2x2 (condition x imitation round) repeated measures ANOVA on imitation scores revealed a statistically significant interaction, F (1, 30) = 8.9, p = .006, partial η^2 = .23. Planned comparisons revealed that imitation scores in the *foreign + same* condition decreased significantly from the first (M = 1.88, SD = .88) to the second (M = .56, SD =.89) imitation round, t (15) = 4.61, p < .001, Cohen's d = 1.15; Wilcoxon signed ranks test, Z = -2.91, p = .002. However, imitation scores in the *same + foreign* condition did not change significantly from the first (M = 1.63, SD = 1.2) to the second (M = 1.38, SD = 1.15) imitation round, t(15) = 1.17, p = .26, Cohen's d = .29; Wilcoxon signed ranks test, Z = -1.13, p = .5.

Demonstrations from the same- and the foreign-language speaker were attended to equally well (both medians = 100% of the demonstration duration).

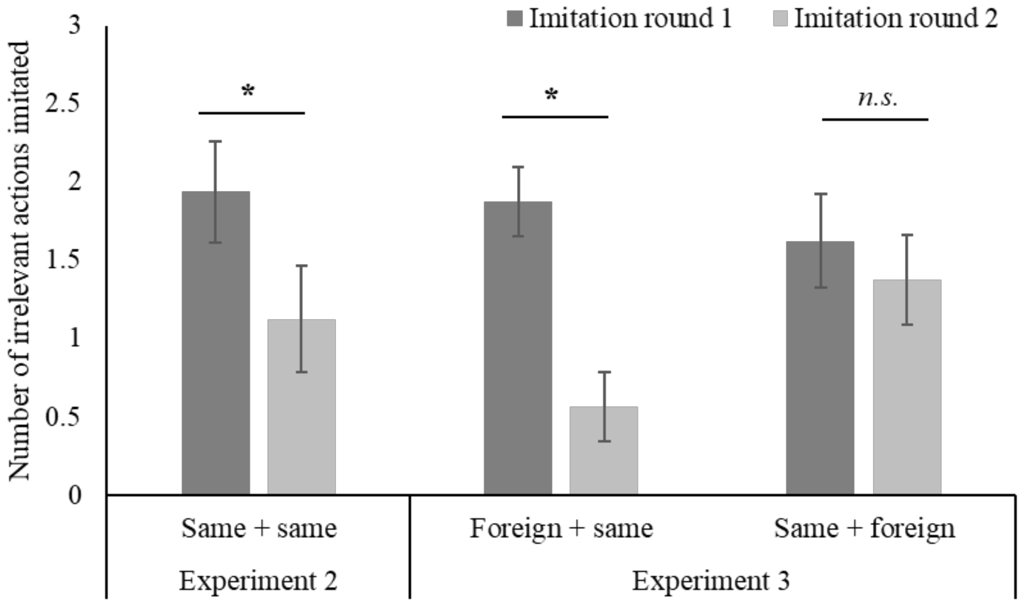

Figure 5.3. Mean number of irrelevant actions imitated by four-year-olds in each imitation round in experiment 2 (*same* + *same*) and in in two conditions of experiment 3 (*foreign* + *same* versus *same* + *foreign*). The asterisk marks a statistically significant difference, *n.s.* marks non-significant difference. Error bars show ±1 *SE*.

We further carried out a 2x2x3 (condition x imitation round x age group) repeated measures ANOVA on imitation scores with the participants of experiment 1 and 3. This indicated no statistically significant three way interaction between these factors, F (2, 90) = 2.38, p = .098, partial η^2 = .05, but a statistically significant interaction between the imitation round and age group, F (2, 90) = 6.7, p = .002, partial η^2 = .13. There were no differences between three age groups in their imitation scores during the first imitation round, F (2, 93) = 1.44, p = .24, partial η^2 = .03; $\chi^2(2)$ = 2.15, p = .35. Yet both parametric and non-parametric tests revealed a statistically significant difference between the age groups in their imitation scores during the second imitation round, F (2, 93) = 3.45, p = .04, partial η^2 = .07; $\chi^2(2)$ = 6.99 .34, p = .03. Post hoc Bonferroni comparisons showed that 4-year-olds' imitation scores were higher in the second round in contrast with 6-year-olds' (p = .04) but not in contrast with 5-year-olds' (p > .05). There were no differences between the imitation scores of 5- and 6-year-olds in the second imitation round (p = .15).

5.4.5 Discussion

Similar to experiment 2, in experiment 3 4-year-olds readily imitated the sub-efficient action routines that they had acquired from the ostensive demonstration of the first model. Notably, this tendency to faithfully imitate the sub-efficient routine was present irrespective of whether it was initially demonstrated to them by a speaker of their own language or a speaker of a foreign language. Furthermore, just like in experiment 2, the 4-year-olds in this experiment reduced their imitation of the superfluous steps of the sub-efficient action sequence after a more efficient alternative was presented to them by a speaker of their own language.

Crucially, consistent with our prediction, these younger children continued to preserve and use the initially acquired sub-efficient action routine, even after the more efficient alternative was subsequently demonstrated to them by the foreign-language speaker. Children's persistence in imitating the sub-efficient means after being exposed to the efficient alternative from a foreign-language speaker could not have stemmed from mere reluctance to imitate the speaker of a foreign language, as illustrated by the high level of imitation in the first round in the *foreign + same* condition. Neither could it be due to inability to modify their imitation of the means acquired from a same-language speaker, as children in experiment 2 were clearly able to do just that. We conclude that 4-year-olds selectively either adopted or disregarded the efficient alternative in experiment 3 as a function of whether it was demonstrated to them by the speaker of their own language or by the speaker of a foreign language. This result is in line with our proposal that in young children the retention of sub-efficient customary ways of doing things may be aided by their selective evaluation of the cultural competence of the source of the efficient alternative. This selectivity depends on whether models can be recognized as members of the cultural community sharing a common body of cultural knowledge the child is growing up in.

5.5 General Discussion

The present chapter investigated how children optimize their acquisition of common cultural practices that are shared by the members of their community. We focused on the learning and retention of sub-efficient action routines because – while they may appear arbitrary and remain cognitively opaque to naïve learners – they are likely candidates for customary and normative culture-specific practices that are part of the common cultural knowledge repertoire shared by members of a community (Legare & Nielsen, 2015). As such, they should be prime targets of cultural transmission and the young naïve cultural learner should be able to acquire and maintain them in their repertoire despite the availability of efficient alternatives (Gergely & Csibra, 2006; Kiraly et al., 2013).

We proposed that from early childhood this process is aided by the tendency to embrace later encountered more efficient alternative means only selectively depending on whether the efficient alternative comes from an agent who is considered to possess the cultural knowledge of the child's community, as evidenced by the language the agent speaks. We hypothesized that having learned a sub-efficient way of achieving a goal, preschoolers should adopt the more efficient alternative if it was demonstrated by a shared-language speaker, but should be more likely to disregard it (and retain the sub-efficient action routine already acquired) if its source is a foreign-language speaker.

To sum up our three experiments: contrary to our hypothesis, after having acquired the sub-efficient action routine demonstrated to them, 5- and 6-year-olds (experiment 1) flexibly adopted the novel but more efficient alternative means action presented by the second model regardless of whether the model was a speaker of the shared or the non-shared language. On the other hand, whether 4-year-olds (experiments 2 & 3) adopted the efficient alternative indeed depended on the language spoken by the person modeling it.

The difference between 4-year-olds and the older age-groups was not predicted and further studies are needed to clarify it. For now, we suggest that 5- and 6-year-old children may have already come to possess a broader scope of contextually relevant knowledge that they could more flexibly integrate in their evaluation of an informer's likely cultural competence. For example, by 5- or 6-years of age preschoolers may have encountered in their social environment sufficient number of relatives, familiar adults or kindergarten teachers, who could speak more than one language but nevertheless proved to be reliable sources of shared cultural knowledge. Given their growing experience with people speaking multiple languages in their cultural environment, these older children may have come to rely less on language spoken as a strong indicator of cultural competence. In particular, they may not consider evidence of speaking a foreign language as sufficient to evaluate the source agent as lacking relevant culture-specific knowledge shared by the child's social community. It should be noted, however, that in this study we did not assess children's exposure to languages in their social environment and thus were not able to test this post-hoc account of the observed differences between age groups.

We also recognize that our procedure itself might have played a role in the developmental trend we found. First of all, we carried out this paradigm in the kindergartens of the children. This alone might have led them to perceive all three female adults (the experimenter and two models) as belonging to the same cultural group, independent of the language the models spoke. More crucially, after an ostensive and detailed explanation of the game-structure to the children during the warm-up phase, the experimenter explicitly called their attention to the fact that they should watch attentively the models who will be coming soon to obtain stickers from the sticker box. This ostensive introduction by itself might have led children to construe the models as knowledgeable about the goal of the game and about how to work the novel apparatus - indicating that they are reliable informants.

Previous research documented similar age differences in the way verbal cues influence children's imitative flexibility, suggesting that as children get older, they might be more attuned to the content of the verbal framing of the task (Clegg & Legare, 2016). In our study, the outcome-oriented language used during the introduction might have led, especially the older children to take an instrumental stance in interpreting the task and to prioritize the efficient attainment of the sticker over the sub-efficient action routine. For 4-year-olds, on the other hand, the verbal content of the communicative instruction about the task context and the implication of the verbal instruction concerning the likely knowledgeability of the upcoming demonstrators about the task structure might have played a lesser role and didn't yet interfere with their language-based evaluation of the model as a likely source of cultural knowledge.

The current results are in line with the view that from early on in their development human children are proficient learners of culturally shared knowledge. Young children form different expectations about the generalizability of the newly acquired information depending on whether it was ostensively communicated to them or not (Bonawitz, Shafto, Gweon, Goodman, Spelke, & Schulz, 2011; Butler, Schmidt, Buergel, & Tomasello; 2015; Csibra & Gergely, 2006; 2009; 2011; Kiraly et al., 2013). They readily acquire novel means actions regardless of their apparent sub-efficiency (Hoehl et al., 2014; Horner & Whiten, 2005; Lyons et al., 2007; 2011; Schleihauf, Graetz, Pauen, & Hoehl, 2018), show selectivity in reproducing sub-efficient action routines depending on their audience (Nielsen & Blank, 2011), and transmit such sub-efficient means actions to third parties (Kenward, 2012; Keupp et al., 2013; 2015). Young children also monitor opportunities for learning culturally shared information by attending to the language spoken by informants (Begus et al., 2016; Buttelmann et al., 2013; Howard et al., 2015; Olah et al., 2014; Shutts et al., 2009). As our study shows, around 4 years of age children are also sensitive to the speaker's language as a cue to shared cultural knowledge and rely on it when deciding whether to incorporate the demonstrated efficient action alternatives in their own cultural knowledge repertoire.

Chapter VI. Discussion

The main question this dissertation raised was how young children represent, acquire and maintain cognitively opaque cultural knowledge forms, such as conventional action routines. While instrumental acts can be acquired through exploration and individual trial-and-error learning, conventional knowledge shared by the cultural group, such as traditions, rituals, or social rules, can only be learned from social sources (Tomasello, 1999a, 1999b; Gergely & Csibra, 2006). Conventional actions are also different from instrumental acts in the sense that in order to interpret them one needs to suspend the requirements of instrumental rationality which is an essential constraint on action interpretation in the domain of goal-directed instrumental agency (Gergely & Jacob, 2012). Throughout my empirical work for this dissertation, I operationalized cognitively opaque conventional actions as sub-efficient action routines, given that they violate the requirements of physical/causal efficiency making them either causally (i.e. as the actions lack a physical/causal explanation due to violating causal efficiency) or functionally opaque (i.e. in so far as it may not be transparent for the learner what ultimate goal they may serve), and that they require extra effort to perform them. Based on previous research showing that from an early age young children show imitative flexibility revealing that they do interpret the behavior they observed either as an instrumental or a conventional act depending on certain social cues (Herrmann, Legare, Harris, & Whitehouse, 2013; Legare & Nielsen, 2015; Legare, Wen, Herrmann, & Whitehouse, 2015), I argued that any purposefully displayed or demonstrated action that cannot be interpreted with the framework of instrumental rationality could trigger a socially motivated interpretation among naïve observers/learners in certain social contexts (cf. ritual stance, Legare, 2019; Legare & Harris, 2016).

My hypothesis was that acquiring ostensively demonstrated conventional actions would provide an epistemic benefit for the naïve learners as they bring the naïve learners' attention to the socially shared and culturally useful new action routines that they would be motivated to master in order to be culturally competent members of their social group. In other words, I predicted that being demonstrated with such actions that do not causally relate to the end goal achievement in the most efficient way available could instantiate an interpretation in the naive observer akin to the reading of "this is how *we* do it among our group". Given the close link between shared language and cultural common ground, I argued that young children will be more likely to learn and retain these sub-efficient action routines when these are demonstrated to them by their linguistic group members. It is also likely such actions also provide further benefit as they can act as informative signals indicating one's commitment or belongingness to her social group by the virtue of acting as credibility enhancing displays (Henrich, 2009). In this reading, I proposed that shared cultural repertoire between agents could be a cue revealing that they belong to the same social group, guiding naïve learners' predictions and inferences about what kind of actions agents from the same group *ought to* perform in certain social contexts.

Below, I first review the findings of the empirical chapters I have incorporated into this dissertation, along with how I think the empirical work presented in each chapter may also give rise to further interesting questions to explore in future work. In the remaining sub-sections of the discussion I aim to lay out how my empirical work feeds into the existing theoretical work on natural pedagogy, action understanding and norm psychology.

6.1 Summary of the findings and future research

In Chapter 2, I tested 18-month-old infants in a head-touch imitation paradigm. This study was set out to investigate whether infants selectively learn ostensively demonstrated sub-efficient manner of actions from their linguistic ingroup members. The main finding was infants' selective imitation of sub-efficient demonstration as a function of the language the demonstrators spoke. Eighteen-month-old infants reenacted sub-efficient actions that was demonstrated to them by a speaker of their own language in contrast to a speaker of a foreign language. Additional analyses showed that infants, despite emitting the light effect with their hands first, still reenacted the head-touch action as shown by same-language speaker strengthening the interpretation that they did not only aim their behavior to be in line with the demonstrator's in terms of the visible outcome it achieved but further on the exact way the goal was achieved . On the other hand, infants in the foreign-language condition seemed to persevere with the first action they performed (the efficient hand-action) as long as it successfully activated the light box. Note that, imitation behavior that was elicited from the infants was not prompted by any verbal instruction and neither the respective demonstrator nor any other experimenter was present in the room during infants' response period. I argued that the infants did not consider the only goal of the demonstration as achieving the light effect when the demonstration was delivered by the informant speaking the same language with them, but also took into account of the specific way that end goal was achieved. This study, therefore, did not only replicate Buttelmann et al.'s (2013) findings of the effect of the same-language with a slightly older age group, but also extended the implications of the results by incorporating additional measures such as whether the head-touch actions performed by the infants were preceded by their hand actions first, which already enabled them to illuminate the box in a more efficient manner.

Still, I believe it will be an interesting avenue for the future research to investigate whether infants would reenact the sub-efficient means displays of their own social group members even after they have initially realized the visible goal by performing a more efficient means actions themselves first, before the demonstration phase. This design would also be informative as it would show whether infants can inhibit the efficient mean actions (i.e. hand action), that is driven by the experienced affordance properties of the light box, while later adopting the demonstrated sub-efficient action. For example, Pinkham and Jaswal (2011) found that when 18-month-old infants were given the opportunity to explore and efficiently realize a goal by themselves first, they did not modify this response by faithfully imitating a sub-efficient action manners that was demonstrated to them later. Note that in this study the experimenter who handed the light-box to the infants after her ostensive head-touch demonstration prompted the infants to act on the box by asking them “Can you show me how it works?” and remained in the room sitting opposite of the participants throughout the response period (as if to verify whether children can effectively obtain the end result as instructed.) Whether 18-month-old infants’ selective imitation in this study was influenced by what they understood from this verbal instruction, instead of the experimental manipulation presented by the researchers, remains unclear and would present a further empirical question to be pursued by future research. Yet there is also evidence documenting that 14-month-olds can, in fact, take into account the verbal framing of the demonstrated act and imitate accordingly (Chen & Waxman, 2013). Given this, it is not clear whether infants would be willing to reenact the sub-efficient mean action even after efficiently realizing the goal itself in a context in which they are not prompted by any verbal directives. I would predict that they would be imitating the sub-efficient action demonstration when the instruction does not highlight the effective method of illuminating the box.

The findings of Chapter 2 further strengthen the evidence that cultural learning takes place from early on in ontogeny. During their second year, infants go beyond the constraints of their teleological action interpretation system that they had readily relied on from the age of 3 months and incorporate seemingly sub-efficient and cognitively opaque action means such as a head-touch action into their action repertoires despite the availability of alternative efficient means actions. This suggests that they can also organize the observed event sequence hierarchically from a rather early age: when the mean action does not relate to the goal achievement in the most efficient way available, they show readiness to encode the mean action if it is ostensively demonstrated to them by their own language speaker as an essential component of the demonstrated event sequence that is to be retained as relevant. Probing the relative strength of this mean action representation is another relevant direction to be pursued by future research that I hope the current work will give a rise to. If I am right in my interpretation that infants construe sub-efficient mean actions as relevant sub-goals, then encoding such sub-goals would allow infants to make predictions about how the hierarchically higher goals *should be* achieved among the same social group members. In Chapter 3, I aimed to test this hypothesis. However, admittedly, one could have designed a significantly simpler looking time study than the one presented in Chapter 3 to tap into the same question. A recent study by Liberman and her colleagues (2018) is a good example of this possibility.

My empirical investigation in Chapter 3 to explore whether infants can expect a particular causally opaque action form to be shared only among the members of the same social group did not yield a clearly interpretable finding. The rationale for this study was based on the assumption that imitation of sub-efficient means is in part driven by the fact that such sub-efficient mean actions carry socially relevant information. Specifically, my hypothesis was that sub-efficient means actions may be encoded as culturally "normative" sub-goals when ostensively demonstrated to infants by the speakers of their own language, specifying how an action ought to be performed by the members of the same linguistic group. Infants were presented with an ostensive informant speaking their own language who demonstrated how to operate a touch-sensitive box by leaning forward to contact the box with her forehead. Later they viewed two news agents either speaking in the same language of the informant (so also the same language as their own) or speaking a foreign language. Then these two agents acted on the box by using two novel sub-efficient mean actions. One agent operated the box with her elbow and the other used her head. The agent's head touch action was similar to the ostensive informant’s head action only in so far that both achieved a contact between the surface of the box and the forehead, but the agent lifted the box towards her forehead instead of leaning on it. Assuming the functional equivalence of these two emulative variants of achieving contact between the box and the head, I predicted that infants would look longer at the more dissimilar elbow action in contrast with the lift-to-head action when the agents performing them belonged to the same linguistic group as the ostensive informant. Instead, the findings with 14-month-olds revealed longer looking times to lift-to-head action. A follow-up experiment with 18-month-olds, in addition to post-hoc designed baseline conditions, suggested that 14-month-olds’ longer looking times to one action over another were likely to stem from the relatively higher salience of the lift-to-head action in comparison to the elbow action. It was the low-level featural differences of the test actions guiding infants’ differential visual attention to the test events, and not their evaluation of the functional similarity the emulative versions of achieving the sub-goal (make contact between hand and box) that the lift-to-head action and ostensively demonstrated “leaning-over-box-with-the-head” actions represented. It remains as an open question whether infants expect same group members to adopt the same sub-goals when tested in more carefully designed looking time paradigms.

Future research could investigate whether infants would expect common cultural ground to be shared among those who speak the same language as in the studies by Liberman and her colleagues (2018) and Olah and her colleagues (2014). Instead of probing infants’ expectations on the kind of causally opaque actions the agents are likely to perform, as was done in Chapter 3, future work could explore infants’ expectations about whether two agents would speak the same language or not after infants had been familiarized with the same or different sub-efficient mean actions performed by the two agents. Specifically, one can predict that if two agents perform the same sub-efficient mean action in bringing about the same goal despite the availability of an alternative more efficient mean action, they would be expected to speak the same language later. Likewise, if the two agents had been previously seen as acting on a light box by using the same sub-efficient means but this time this would be motivated by a relevant physical constraint, infants should not expect them to speak the same language. Given the complex stimulus design that was used in Chapter 3, in Chapter 4 I relied on infants' early capacity to classify agents as belonging to the same group based on cues such as their similar physical appearance, their spatial proximity, and their synchronous movements building on the work of Powell and Spelke (2013) to investigate infants' expectation of common behavioral repertoire among the members of the same social group.

In Chapter 4, I tested 11-month-old infants in a violation of expectation looking time paradigm. This study was set out not only to investigate whether infants expect group members to act alike, but also gave us a platform to examine whether and how this expectation of shared-movement-repertoire between group members interact with infants' expectations about efficiency of goal-directed actions. For this research I relied on infants' early social categorization based on cues indicating similarity (i.e. color, shape) and interaction patterns (i.e. proximity, synchronous dancing) between animated characters. Across a series of experiments, infants viewed two agents from the same group moving above a barrier to reach a target object either inefficiently (the barrier was short, allowing for direct approach the goal object, Experiment 2) or efficiently (the barrier was long, hence the only way for the agents to obtain the goal object was to move above the barrier, Experiment 3) during the familiarization. At test, infants showed no clear expectations about the third group member's actions. However, when the test agent did not belong to the same group as the familiarization agents who were viewed as moving above the barrier efficiently in the familiarization, then infants expected efficient goal approach from the outgroup test agent (Experiment 4).

This pattern of findings suggests that infants expect goal-directed agents both to act efficiently to reach the goal, but they also expect them to act similarly as their group members, and that these two expectations may contradict each other (Experiments 2-3). However, when group membership is irrelevant (Experiments 1a, 1b and 4) infants expect efficient goal approach from the test agents. It remains an interesting question at what point in development these two principles of shared-movement-repertoire on the one hand, and rationality of goal approach on the other, could be flexibly used and integrated by young children in helping them to interpret and predict the behaviors of other agents. For example, could older infants expect the novel agents to move alike their own group members despite the inefficiency of the actions performed by their group members previously? Would they govern their predictions of the novel group members based on efficiency, disregarding the movement patterns the same group members had displayed earlier? With this empirical work, I showed that 11-month-old infants' expectation for efficient action was influenced by the shared movement repertoire between social group members. However, the shared movement repertoire between group members did not completely override infants' expectation of efficiency, as revealed by the null findings indicating that both expectations were at work. It is an exciting future avenue to probe how and when young children come to integrate and represent actions flexibly as governed by both social and instrumental principles depending on the particular context.

In Chapter 5 given the recent studies on over-imitation suggesting that children could interpret sub-efficient action routines that are embedded in the demonstrated sequence as conventionally necessary (Nielsen, 2018; Nielsen, Kapitany, & Elkins, 2015) and as normative elucidating on how the activity ought to be performed (Kenward, 2012; Kenward, Karlsson, & Persson, 2011; Keupp, Behne, & Rakoczy, 2013), I used an over-imitation task with older children. I hypothesized that children may rely on their sensitivity to differentiate speakers of their own versus a foreign language as an informative cue indicating whether the demonstrator belongs to their own cultural community and the sub-efficient action routine modeled by them represents shared cultural knowledge. I assessed preschoolers' imitation following two different demonstrations: the first model demonstrated a *sub-efficient action* sequence while the second presented a *more efficient alternative* action to obtain the same goal, by using a paradigm developed by Hoehl and her colleagues (2014). I varied whether the children had heard the models speak their own or a foreign language before their non-verbal action demonstrations. Basically, the models spoke on the phone with an unseen interlocutor either in the same language of the children or in a foreign language before their action demonstrations. Four-year-olds adopted the second model's efficient alternative, but only when she spoke their own language. However, they disregarded the efficient alternative if it was presented by a foreign language speaker and continued to perform the sub-efficient routine they had initially acquired. Four-year-olds, therefore, employed the cue of shared language to optimize acquiring and maintaining culturally shared sub-efficient action routines by selectively updating their action repertoire depending on their language-based evaluation of the demonstrator's culture-specific competence. In contrast, 5- and 6-year-olds adopted the efficient alternative independently of the demonstrator's language.

I argued that children's growing experience with multilingual speakers in their social environment may have made it already easily acceptable for older children that an over-heard phone-conversation in a foreign language does not rule out the possibility that the person is *also* a speaker of the child's own language. In a related vein, the verbal framing of the demonstration by the experimenter (who ostensively announced that "there are going to be people coming to get a sticker out of this box") might have sanctioned the foreign-language speaking model as being knowledgeable about the apparatus thereby providing reason for older children to take an *instrumental stance* when interpreting the task demonstration (see Clegg & Legare, 2016, showing that as children get older they are more attuned to the verbal framing of the task). It will be interesting to explore whether children could adopt sub-efficient action routines selectively from the informants speaking the same language with them after having observed the efficient goal attainment first. Viewing a demonstrator opting for the sub-efficient action routine after the demonstration of the efficient goal retrieval might help children to have a more conventional reading of the sub-efficient action routine demonstrations.

Another line of future research should address the methodological confounds I outlined in the discussion section of Chapter 5. For example, a recent study by Buttelmann and his colleagues (Krieger, Aschersleben, Sommerfeld, & Buttelmann, 2020) used video demonstrations to eliminate ostensive introduction of the demonstrators and avoided verbally framing of the task that emphasized the outcome of the demonstration in their investigation of whether 6-year-old children would adopt an efficient action alternative selectively from the informants sharing the same social group with them. In their study the informants did not speak before, during or after their respective demonstrations, and the only cue that was available to children that could indicate social group of the informants was their physical appearance (i.e. race, Asian or Caucasian). However, it is not clear whether younger children could encode race spontaneously as older preschoolers do to guide their learning from others (for the evidence indicating that they do not, see Diesendruck & HaLevi, 2006; Kinzler, Shutts, DeJesus, & Spelke, 2009; Krieger, Möller, Zmyj, & Aschersleben, 2016; Shutts, Banaji, & Spelke, 2010, and also see why race is not a dedicated dimension of social categorization, Kurzban, Tooby, & Cosmides, 2001). Future research might pit different dimensions of social categorization against one another in different age groups with the aim to present a coherent developmental story on how and why such cues influence children’s learning from different informants.

The research presented in this dissertation has built on different methodologies across different age groups to investigate how sub-efficient action routines are represented and acquired. To sum up my empirical journey outlined across four chapters, beginning from the age of 11-months infants show sensitivity to the shared-action repertoires of social groups. Before their second birthday they readily acquire sub-efficient routines from the informants speaking in the same language with them replicating the previous work of Buttelmann and his colleagues (2013). During preschool years they also rely on the informants' language as a cue of shared cultural ground and incorporate efficient alternatives into their knowledge repertoire, depending on whether the informant speaks the same language as them.

6.2 Actions in context: assumption of relevance

Our early emerging propensity to reason about the observed behavior of others with the presumption of function (Csibra & Gergely, 2007) helps us to attach a purpose to even seemingly "purposeless" actions, helping us to adopt a conventional (or ritual) stance (Legare, 2019). There are indeed several cognitive biases that could enable young children to go beyond determining the physical goals of other's object-directed actions when those object-directed actions appear to violate the principle of rationality. One bias is our propensity to be responsive to pedagogical intentions of others around us. I argued that another bias that works in tandem, is our propensity to differentially attend to and learn from the speakers of our own language.

According to natural pedagogy theory, infants display sensitivity to ostensive signals (such as infant directed speech or eye contact) within days after birth, and such signals help them to recognize the communicative intention of others (Csibra, 2010). Not only do these signals act as attention-enhancing displays for infants leading them to simply orient towards the source of these signals, but they generate an expectation of relevance in infants (Csibra, 2010; presumption of relevance, Sperber & Wilson, 1987). In addition to invoking search for the relevant and new information, ostensive signals are also argued to help their addressees to take an inductive leap from a single observation facilitating the representation of shared knowledge that is applicable and available to other members of one's cultural group (Csibra & Gergely, 2006, 2009, 2010; Egyed, Kiraly, & Gergely, 2013; Butler, Schmidt, Burgel, & Tomasello, 2015). The role of ostensive signals in licensing naïve learners to construe the information demonstrated as relevant to acquire, which in turn guides their learning about culturally opaque knowledge forms, has been demonstrated by previous research (Kiraly, 2009a, 2009b; Kiraly et al., 2013; Brugger et al., 2007). Even when the demonstrated action appears to be not the most efficient way to achieve the physical goal the expectation of relevance induced by ostensive signals enable infants to construe the demonstrated action as manifesting new and relevant knowledge that is worth acquiring (Gergely & Jacob, 2012).

However, as I documented, naïve learners also take into account whether or not the informant speaks the same language as them when interpreting the relevance of the demonstrated action routine. Given the early orienting response of infants to the informants speaking the same language with them (Mehler, Jusczyk, Lambertz, Halsted, Bertoncini, & Amiel-Tison, 1988; Moon, Cooper, & Fifer, 1993) one possibility is that shared-language cue also acts as an ostensive signal. Indeed given the co-occurrence of ostensive signals (e.g. eye-contact, motherese, contingent reactivity) with the same-language cue displayed by others in daily environments of infants, it might be plausible to assume that the same-language could acquire the status of an ostensive signal for infants living in monolingual households. After all, same-language cue can "unambiguously specify that the infant is the addressee of a communicative act", can be discriminated by young infants and induce a preferential orientation towards their source just like other ostensive signals do according to the criteria that apply to ostensive signals as outlined by Csibra (2010, p.144).

While there should be no doubt that the cue of shared-language can code for communicative intention, my findings suggest that the shared-language cue equips the infants with a differential inferential interpretation capacity over and beyond the presumption of relevance induced by ostensive signals. Shared-language cue is a more direct cue of relevance signaling that the informative content of the communicative intention is based on the shared cultural common ground between the teacher and the addressee. Hereby, I have no means to entertain that 18-month-olds have a full-fledged understanding of what culture or cultural relevance could be, yet I argue that infants have an understanding of how they ought to do something and acquire this basic conventional rule selectively from the experts of the cultural group they are socializing into. I also believe that it will be an interesting research endeavor to explore whether same language cue could invoke biases in naïve learners that enable them to form generic and enduring representations about the content of the knowledge transmitted by others who speak the same way as they do. Findings of Chapter 3 failed to show evidence in favor of this generalization effect. Yet given the identified limitations of the methodology, it might be still worth pursuing a similar line of research to probe whether same-language cue could invoke similar generic expectations in infants about the information content as ostensive signals do (Egyed, Kiraly, & Gergely, 2013). In a similar vein, it is an avenue for future research to pit ostensive signals against shared-language cue to investigate for example, whether an ostensive foreign-language speaker would be perceived as a more relevant epistemic source when it comes to cognitively opaque knowledge forms in comparison to a non-ostensive but a native-language speaker. If I am right in my claim that same-linguistic cues have a special role in infants and young children's learning and attribution of cultural forms of knowledge, then one should find the influence of same-language cue in learning, albeit perhaps to a lesser extent, in non-ostensive demonstration contexts, especially when the informative content is novel.

Yet just to note, it is methodologically challenging to decouple the event where an informant reveals what language she or he is speaking from the event of an action demonstration. It is likely that if the demonstrator speaks to the participant just to give the evidence that she is a native speaker of the language of the participant, even if the subsequently shown event by the same agent is stripped down from ostensive signals, participants could still attribute communicative intention to the agent. It was on purpose that these two events were decoupled from each other in my work. Both in Chapter 2 and in Chapter 5 infants and children were given the evidence that the demonstrator either spoke the same or a different language as themselves before the demonstration phase. In Chapter 2 this was done in a story-telling event where both demonstrators socially engaged with the infant. In Chapter 5 this was done in a pretend-phone-conversation the demonstrators were viewed as having. Critically, in the demonstration phases, both informants used the same ostensive signals, such as establishing eye contact. In a different context where the demonstrators are socially aloof both during language induction and the demonstration phase, would participants still selectively imitate the experimenter who gave indirect evidence that she is the speaker of the infants' own language? Previous research (Kiraly et al., 2013) documented that infants do not adopt sub-efficient mean action displays to the same degree when these are staged in non-ostensive contexts. If infants were overhearing the agent speaking with someone else in their own language or in a different language, and then acting on the light-box without any display of ostensive signals, perhaps then infants would not show any difference in their imitative behaviors as a function of the language spoken by the experimenter. In this account, ostensive signals are crucial to infants' interpretation of the relevance of the observed behavior. Yet given the display of ostensive signals by the speakers of their mother-tongue in their daily experience, perhaps infants could also conceive any instance of verbal acts, as long as they are produced in the language they are learning, as acts of ostension. I believe there is a value in probing whether these two cues of ostension and same language interact in guiding infants' learning of cognitively opaque knowledge forms from social models.

6.3 From motor resonance to like-me hypothesis and to mentalistic accounts

While in this dissertation, the focus of some of my work was on how children imitate in different social contexts I did not capitalize on bottom-up theories explaining the mechanism how infants and young children can imitate in the first place. Proponents of motor resonance theory argued that observing others' actions facilitates performance (Paulus, Hunnius, Vissers & Bekkering, 2011a; Rizzolatti & Craighero, 2004) by eliciting motor simulation in the observers when the action is already in the observers' motor repertoire. These motoric simulations activate low level motor areas corresponding to the observed actions, and then this activation is thought to propagate in a bottom-up manner revealing the potential goals (or intentions) which would have given rise to them. The mirror neuron system has been suggested to constitute the neural basis of this capacity, documenting in several studies (Buccino et al., 2001; Gazzola & Keysers, 2009) that action perception and action expectation invoke activation from overlapping brain areas. This was further taken as evidence that perceiving others' actions and executing those acts have an overlapping representational format (Prinz, 1997).

According to this approach, in order to be able to imitate infants need to match the observed action on their own motor repertoire. Only when the motor system of the observer resonates with the observed action, the observer can understand the outcome the action achieves. Automatic process of motor resonance is argued not only to enable infants to bridge the gap between others' actions and their own action repertoire, by acting as a simulation device, but also to mediate observational learning about action-effect contingencies. According to this model of imitative learning (two-staged model: Paulus, Hunnius, Vissers, & Bekkering, 2011a; Paulus, Hunnius, Vissers, & Bekkering, 2011b) observing another person's action activates the corresponding motor code in infants' motor system if the observed action is already within infants' motor repertoire. If furthermore, the observed action has a salient effect, representation of this effect is associated with the activated motor code forming the basis of bottom-up identification of goals (Elsner & Aschersleben, 2003; Elsner & Hommel; 2001; see also Paulus, 2012 on ideomotor account).

Several behavioral findings showed that infants' own experiences and their motor capabilities are necessary for them to be able to make sense of others' actions and also to perform them (Sommerville & Woodward, 2005a; Hauf, Aschersleben, & Prinz, 2007; van Elk, van Schie, Hunnius, Vesper, & Bekkering, 2008). Given the effect of prior experience on infants' perception and performance of motor acts, motor resonance could be an important factor of action perception from infancy. However, this automatic process postulated by Paulus and his colleagues' two stage model (2011a; 2011b) could not be the only mechanism that subserves infants' imitation behavior and understanding of other people's actions. It is also not clear how this two-stage model of selective imitation can account for the findings of Chapter 2. After all, both demonstrations were the same, hence there should be no difference in motor resonance elicited between the demonstrations of same- and foreign language speakers. Motor resonance accounts also do not adequately explain how infants can make predictions about novel animated characters' action repertoires in highly unfamiliar contexts as shown in Chapter 4 (for related arguments against motor resonance theory see Buttelmann & Zmyj, 2012; Zmyj & Buttelmann, 2014).

As the proponents of two stage models argue other processes might play a role in infant's selective imitation. For example, according to Csibra's (2008b) action reconstruction account, observed actions receive interpretation within the visual system before being transformed into a motor code. What is mapped during mirroring thus is not an "uninterpreted signal but a description of the observed action at some level of the action hierarchy" (Csibra, 2008b, p. 441). While motor resonance (or direct matching) hypothesis posits a mechanism allowing for bottom-up propagation of activation from low-level motor resonance supporting a simulation system that facilitates goal understanding, the action reconstruction account posits top down propagation allowing for goal prediction for observed actions based on how the actions are interpreted. Hence in the latter account, goal understanding is not the output but the input of this mirroring process (Csibra, 2008b).

To my interpretation, action reconstruction introduces an integrative framework for explaining both goal attribution and goal prediction as flexible capacities malleable to contextual factors. This mechanism of action reconstruction would fit our findings better, not only explaining infants' selective imitation of sub-efficient action routines from the informants speaking the same language with them but also how they can form action predictions based on viewing unfamiliar animated agents on the scene before their first birthday as Chapter 4 documented. Relatedly, one explanation for the selective imitation findings of this dissertation is that imitation is modulated by top-down factors (Marshall & Meltzoff, 2014). Yet this is not to say motor resonance has no role in action understanding and in imitation. After all, there is always a cross modal match between the actions when infants see others acting around them, and when they themselves can act like others (Meltzoff, 2007a). This does not only allow infants to motorically relate their own behaviors with observed actions, but also allows them to recognize shared behavioral states, helping them to identify others as "like me".

According to the like-me framework of Meltzoff (2007a, 2007b), recognizing self-other correspondence allows infants to have an interpretative framework to make sense of the behaviors they see. Meltzoff (2007a) suggests that based on recognizing shared behaviors between themselves and others, infants can use their own experiences to understand others' actions, goals, intentions and they also can learn about the consequences of their own actions even before producing those actions by observing the behaviors of others. This bi-directional framework posited by Meltzoff (2007a, 2007b) relies on the tight coupling of perception and production, allowing infants to make inferences and predictions both from self to other and from other to self. This is perhaps more strikingly indexed by studies showing that around the age of 14-months infants are sensitive to those who act like themselves as evidenced by their longer looking and heightened positive affect at an agent who imitated them in contrast to another agent who acted contingently but differently to themselves (Meltzoff, 1990; also see Carpenter, Uebel, & Tomasello, 2013; Over, Carpenter, Spears, & Gattis, 2013). Infants do not seem to only recognize when an agent acts in the same manner as they do, and they actively prefer the social partner who behaves like themselves.[6] Human infants from early on also imitate the acts of others to fulfill the goal that could not be realized in the demonstration but only if the agents performing them are more similar to themselves (human agents vs mechanical device, Meltzoff, 1995).

The like-me hypothesis might seem to fit well with our data in Chapters 2, in explaining infants' tendency to selectively imitate those agents who are like themselves, by the virtue of the same-language they speak with the participants. From this point of view, it appears that like-me preferences could be indeed another factor underlying selective imitation of young children from informants sharing the same social group with them. However, as previous studies documented, not all who are "like me" induces similar biases among young children (Kinzler & Dautel, 2012; Kinzler & Spelke, 2011 on infants' and children' treatment of the racial-group membership cue). Furthermore, guiding our learning by only relying on others who are similar to ourselves would not be an effective mechanism by which we can learn about the world, considering the multitude of other cues which could more directly indicate epistemic reliability of a potential informant. For example, young children rely on adult informants in

[6] Recognizing being-imitated is not human-specific. Paukner and his colleagues (Paukner, Suomi, Visalberghi, & Ferrari, 2009) documented that capuchin monkeys prefer to approach and exchange tokens with the human agent who exactly copied their actions instead of another human agent who contingently reacted to their actions. Haun and Call (2008) also documented imitation recognition in the four specifies of non-human great apes.

comparison with their peers in learning of novel sub-efficient means actions (McGuigan, Makinson, & Whiten, 2011; Wood, Kendal, & Flynn, 2012; Zmyi, Daum, Prinz, & Aschersleben, 2008) or of novel game activities (Rakoczy, Hamann, Warneken, & Tomasello, 2010). In different contexts, they were also found to rely on their peers if their peer models were more knowledgeable than the adult models (when age was pitted against knowledgeability: Jaswal & Neely, 2006; or in the context of toys: VanderBorght & Jaswal, 2009). It seems that young children's evaluation of the information source with respect to her or his competence overrides any *like-me* preference they might have.

Additionally, the like-me framework falls short of explaining the findings we have in Chapter 5. Preschool children were not more likely to imitate causally irrelevant actions in the first imitation round when these actions were modeled by the informant speaking the same language as them in contrast to when these were modeled by a foreign-language speaker. Moreover, older children were likely to adopt the alternative efficient means independently of the language the models were speaking. Thus I argue that the same-language cue does not encourage selective learning among young children only by the virtue of signaling them that the agent speaking the same language is like them (after all, if that was the case any cue which signals that the model is "like-me" should have yielded similar biases in infants and young children), but is interpreted as a rich cue indicating the potential of the agent in providing culturally relevant information. As we discussed at length in Chapter 5, such an interpretation of the same-language cue is malleable to other social cues and contextual factors that could inform the naïve observer about the relevance of the models' demonstrations (i.e. ostensive signals, introduction of the models as competent informants, etc.). Furthermore, the like-me framework alone does not explain how young children can make sense of actions performed by animated agents, puppets, or robots (or how children can learn from them) as they do not necessarily show close resemblance to them.

However, from a different perspective, it might be rather the case that the heuristic of "like-me" simply provides a general framework on how infants rely on their perception of others as intentional agents just like themselves. This enables them to use this recognition to make sense of others' actions and furthermore to simulate the consequences of their own action choices (Meltzoff, 2007a). Meltzoff's like-me framework (2007a, 2007b), just like homophily-based accounts (Haun & Over, 2015) which capitalize on our preference for and reliance on others who are similar to us from early on in ontogeny, could be foundational blocks for our early emerging capacity to construe others' behaviors as governed by unobservable states. For example, to any cognitively mature observer it is clear that an agent who is reaching for food does so because he is hungry. Based on the same behavior one can also see that the agent is reaching for food because it is her favorite food (not because she is hungry), or because it would be rude to say no to the host offering food (in certain cultures). Any observed action can be explained in relation to a goal; and goals can be represented in different levels of the goal hierarchy depending on contextual cues in order for us to explain the behaviors of others who have the same biological, psychological, physical and social constraints on their actions as we do. Recognizing others who are similar to ourselves could be the basis of an understanding that their behaviors too are governed by the similar constraints as ours, even though we cannot always deduct those constraints from the physical environment alone. In this reading, similarity-principle does not lead us to selectively attend to and learn from others who are similar to ourselves but supports our capacity to read intentions and motivations behind the observable states of similar others in general.

The mentalistic account of infant imitation argues for the same intention reading capacity as underlying infants' imitative skills (Buttelmann, Carpenter, Call, & Tomasello, 2008). This account draws on the findings that around their first birthday, infants' understanding of intentional actions extend beyond determining the physical goals of others' object-directed actions (Carpenter, Akhtar, & Tomasello, 1998; Meltzoff, 1995; Tomasello & Barton, 1994). Mentalistic accounts therefore suggest that infants understand that others chose their actions for a reason even in contexts when that reason cannot be directly perceptually observable. I argue my findings in imitation studies are in line with this account. Not only do infants reason about the rationality of the action in the way the action causally relates to the goal from the perceptual input alone, but moreover they can entertain that the behaviors of others are guided by intentions and reasons that cannot be readily observable. Early sensitivity to ostensive cues indeed could boost this intentional reading by calling the naïve leaner's attention to the fact that the demonstration episode is addressed to and tailored for them, inducing basic trust in the source of this communicative intention. Same language spoken by the model further reveals the cultural relevance of the information provided to them, further acting as a cue informing the learner about the relevant epistemic state of the informant about cultural-specific behaviors that she or he is expected to have a competence of. Recognizing communicative intent and interpreting the content of this intent being about the cultural common ground, in which the naïve learner is in the process of participating, could only be possible if the learner could entertain that agents act for reasons.

6.3.1 Lessons from autism spectrum disorder

Difficulties with social and communicative behaviors are diagnostic characteristics of autism spectrum disorder (American Psychiatric Association, 2013). Hobson and Hobson (2008) (also see Hobson & Lee, 1999) documented that while children with autism are proficient in motor control and can imitate goal-directed actions of others, when it comes to imitating the distinctive style the demonstrator adopted while performing those actions, especially when the style was not causally related to the goal retrieval, children with autism were less likely to copy the style of the action demonstration in comparison with the age-matched control group. Further evidence on this comes from over-imitation studies, documenting the lower rates of imitation of children with autism in the demonstrated sequence of causally irrelevant actions in comparison with the performance of typically developing children on objects that are familiar to both samples (Marsh, Pearson, Ropar, & Hamilton, 2013 but also see Nielsen, Slaughter, & Dissanayake, 2013). Children with autism are also found to be performing worse than children with typical development on tasks that tap into their imitation of arbitrary actions and gestures (Wild, Poliakoff, Jerrison, & Gowen, 2012; for a review see, Williams, Whiten, & Singh, 2004).

These findings have been interpreted in the literature as evidence that children with autism have difficulties reproducing the goals that are less salient in the task. This does not only support the view that children with autism lack an understanding of others' communicative intention behind their actions (see Carpenter, 2006), but also suggests that imitation is rather guided by a top-down process, mostly based on social factors (Wang & Hamilton, 2012). These findings coming from children with autism also supports the view that acquiring cognitively opaque action routines from pedagogical informants is a rich and insightful process that is subserved by the observer's capacity to entertain the reason why the actions are performed that way.

6.4 From social category cues to person judgements: a developmental interplay

This dissertation documented the role of social factors in facilitating the understanding of causally opaque actions, showing that infants and children go beyond the efficiency constraint and imitate and/or retain the demonstrated information selectively as a function of the same-language spoken by the model. I suggested that our early developing propensity to form strong representations for social categories enables us to track informants who could be culturally credible sources to rely on while acquiring cultural knowledge of the social group we live in. My proposal was that forming social categories might rather equip us with probabilistic power in inferring others' histories and predicting the likelihood of having common cultural ground with them. This reading implies that the same-language cue is not represented in an either-or manner, leading to categorical person judgements. Instead children, especially as they get older, might have more nuanced reading of the same-language cue, and they treat it as an informative signal with which they can guide their evaluations about the cultural knowledge repertoire of informants. Yet I believe younger children and infants are more likely to treat same- and foreign-language cue as a heuristic guiding them to form categorical representations of others as belonging to "us" and "them". As they get older the same cue becomes just another signal for the likelihood of an individual's possession of shared cultural knowledge repertoire, leading older children to generate epistemically-motivated flexible judgements, which will be malleable to other contextual factors.

Indeed, 18-month-olds in Chapter 2 readily imitated the same-language speaker informant's sub-efficient means in contrast with the foreign-language speaker. On the other hand, preschoolers did not show such a selective imitation behavior. Chapter 5 documented that when older children viewed the sub-efficient action routine either demonstrated by the same-language speaker or a foreign language speaker, 4-, 5- and 6-year-olds all were likely to adopt this demonstration in the first imitation round independently of the same or foreign language the demonstrators spoke. I believe this seemingly discrepant findings that my two chapters revealed on selective imitation could be resolved if the reader considers the argument that the same-language cue is represented differently across age groups.

While the same-language spoken by the demonstrator induced a presumption of relevance among 18-month-old-infants, readily biasing them to adopt the mean action ostensively demonstrated by the informant, preschool aged children had a more nuanced reading of the same language cue given their more sophisticated pragmatic reading of the contextual cues. Yet it is not clear whether the use of ostensive signals was a factor making older children to construe a foreign-language speaker as an equally acceptable source of cognitively opaque mean actions as a native-language speaker. This alone warrants more detailed empirical inquiry so does my interpretation that children' reliance on social group membership cues in learning becomes rather more nuanced as they get older.

Nevertheless, one can claim that the differences in the pattern of selective imitation between age groups (infants and preschoolers) could be the function of the methodological differences of two studies. In Chapter 2, during the phase where infants were given the evidence what language the demonstrators spoke both demonstrators were socially engaging with infants. However, in Chapter 5 children viewed the demonstrators speaking not with themselves, but with an unseen interlocutor on the phone. This implicit induction of the language the informants speak could be one reason for finding no selective imitation in preschoolers (but see Krieger, Ashersleben, Sommerfeld, & Buttelmann, 2020 for replication of no-selective imitation in the first imitation round in a very similar over-imitation task). Future research could probe the developmental trajectory in the extent that infants and children would rely on their linguistic group members in learning causally opaque action routines from ostensive informants.

6.4.1 How privileged is the case of linguistic groups?

The mere fact that minimal groups do not induce same biases among children before the age of 3 (Richter, Over, & Dunham, 2016) and perceptually salient social group membership cues such as same race (Shutts, Roben, & Spelke, 2013; Kinzler & Spelke, 2011) do not seem to generate robust social preferences among young children, suggests that certain cues have a privileged status in shaping children' emerging representations for social categories. In line with previous research (Kinzler & Spelke, 2011; Kinzler, Dupoux, & Spelke, 2007; Liberman, Woodward, & Kinzler, 2017) I argued that infants are predisposed to recognize differences in linguistic repertoires between others from early on as language served as a reliable marker for social alliances throughout the hominid evolution. Note that acquiring native language proficiency is also hypothesized to be constrained by a sensitive (or critical) learning period (for a review, please see Cohen, 2012), which strengthens the evidence that same-language (and same accent) are honest signals informing the observer that the speaker is either born into that particular cultural community where that language is being spoken or had massive extensive experience interacting with those from that cultural community. Given these arguments, it is no surprise that language is treated as a stable and reliable dimension of social categorization from early on (Liberman, Woodward, & Kinzler, 2017). Yet the literature might benefit more from pitting different social group cues against one another to probe naïve juveniles' representation of not only how such cues constrain people's interactions with one another, but also their social behaviors in general.

I provided evidence that infants and children are sensitive to the same-language cue. Yet I did not directly test the hypothesized privileged status of language-cues in none of my studies. For example, would infants be more likely to acquire a sub-efficient action routine from a same-race but a foreign-language speaker in contrast to a different race but a same-language speaker? Even though my answer would be no, this is clearly an empirical question that warrants further research that would also have the value of showing that the same-language cue is not treated differentially just based on the principle of familiarity among young children. Furthermore, it would be interesting to directly pit different social group membership cues to one another in order to investigate the specialized treatment that the same- and foreign-language cues receive from infants from early age on in guiding their inferences about the shared cultural knowledge ground with others as suggested by the work of Soley and Aldan (2020).

6.5 Learning and social functions of imitation

My findings with imitation paradigms in Chapter 2 and 5 could be interpreted as adding another puzzling finding to the literature showing children's selective imitation on one hand, and their *non-selective* faithful imitation on the other – for the first imitation round with preschoolers (for a review see Over & Carpenter, 2012; 2013). Previous research with different paradigms also documented a similar pattern of findings: infants selectively imitate, and older children are more likely to show faithful imitation (and faithful copying seems to increase with age, see McGuigan, Makinson, & Whiten, 2011). These findings, albeit coming from different paradigms, were taken as evidence that infants act on their epistemic motives initially, and only as they get older, do they use imitation socially (Nielsen, 2006; Whiten, McGuigan, Marshall-Pescini, & Hopper, 2009). In other words, it was argued that young children have a more instrumental reading of the task but as they get older, they construe the task as demanding a more conventional interpretation (see Legare, 2019 for a recent review), probably due to their maturing understanding of normativity. In the discussion of Chapter 5, I elaborated on how contextual cues that I did not even aim to manipulate (e.g. outcome-oriented framing of the task, the introduction of the demonstrators as competent agents) could shift children's interpretation of the task as being guided by instrumental principles. These discussions hopefully document the need for future research to take into account rich pragmatic inferences children can bring into situation allowing them to selectively imitate and faithfully produce the action routines they were communicately shown.

However, at this point I should note, while I aimed to test for the epistemic functions of the same-language cue by investigating participants' learning of sub-efficient action routines in the absence of the demonstrators, the studies presented especially in Chapter 2 and 5 do not entirely preclude the possibility that participants imitated for social reasons too. Even when the models were not in the room during the reenactment phase in both studies, participants might have opted to imitate the exact same manner of action that the demonstrators performed just to be like the demonstrator belonging to their linguistic group. As Over and Carpenter (2012, 2013) portrayed in their integrative account on imitation, children themselves could have different motivations while deciding what and from whom to imitate. Building on their theoretical account, I would also argue that social and epistemic motivations could work in tandem when it comes to learning culturally relevant information from others. Yet I argue that epistemic motivation precedes social functions of imitation in these tasks. After all, naïve learners should be epistemically motivated to be competent cultural members, so that they can also signal their belongingness and social affiliation to their own group members later.

Furthermore, while the focus of my studies on infants' and young children's learning and preservation of cognitively opaque mean actions, I did not provide any evidence for whether infants and young children can use what they learned in these tasks communicatively and selectively with other parties. There is already evidence that such epistemically driven acts mark social affiliative relations even for 16-month-olds infants (Liberman et al., 2018), so it would be no surprise that young children could perform sub-efficient action routines just to show their belongingness to their own cultural groups (see Wen, Willard, Caughy, & Legare, in press for related evidence on this). For instance, to probe this, one can have a novel experimenter present in the reenactment phase either belonging to the same group or a different group with participants. This would enable us to explore whether contextual social pressures in group level could influence how children produce the sub-efficient action routines they acquired from a linguistic ingroup individual. In a similar vein, if young children can go beyond the efficient goal attainments to show to their naïve linguistic group members at the same age as them the sub-efficient action routine they acquired from a culturally reliable informant (i.e. an ingroup adult) selectively too, then we can perhaps have a better understanding of the motivation behind their imitative flexibility.

Yet this does not mean social motivations played no role in the experiments documented in this dissertation. Despite having no demonstrator present when it was the participant's turn to act, as Over and Carpenter (2012) argued children might had still aimed to identify on a group level by making themselves more similar to the group in general. Previous research already showed that in certain social contexts, for example when primed with ostracism (Over & Carpenter, 2009; Watson-Jones, Legare, Whitehouse, & Clegg, 2014), preschool-aged children more closely copy the irrelevant actions the demonstrator displayed in bringing about an instrumental goal in comparison to another group that was not primed with ostracism. For sure imitation serves social functions, but I argue in this context of learning cognitively opaque cultural forms as presented in this dissertation young children's imitation to copy others or retain information from those who speak the same language as them could not be explained solely by their need to belong. My imitation studies show that children could prioritize the manner in which the goal achieved when the demonstration was delivered by their linguistic ingroups in certain contexts. Yet this does not mean that they would not use this information later for social-affiliative reasons.

6.6 From imitation to norm psychology

Infants are not only sensitive to being imitated by others (Meltzoff 1990; Agnetta & Rochat, 2004) and they are not only prosocially responsive towards those who mimicked them (Carpenter, Uebel, & Tomasello, 2013), but around the age of 4-months they also show sensitivity to those who imitate one another in third-party observational contexts (Powell & Spelke, 2018a, 2018b). For example, in a series of studies Powell and Spelke (2018a) found that infants expect an agent to approach the group of characters she previously imitated (i.e. sound or motion imitation) in contrast with another group of characters she did not imitate. Follow-up studies (2018a, 2018b) revealed that infants expect the same affiliative relation when the target of imitation was a lone individual rather than a group. Notably, it also seems that this expectation is asymmetrical (2018a): when infants viewed scenes in which this time the group of characters approached the agent who had previously imitated them (in contrast with scenes where the group of characters whom the agent did not previously engage in imitative interaction with approached the agent) infants did not generate any predictions for the social relations between them. In their follow-up work Powell and Spelke (2018b) further showed that pre-verbal infants are not only sensitive to imitative interactions between third parties, but they also prefer imitators over non-imitators.

Infants from very early age on seem to rely on imitative interactions between third parties in guiding their inferences about the social affiliative relations between these parties. Previous work from the same authors (Powell & Spelke, 2013) further revealed that infants also expect individuals belonging to the same minimal group to mimic each other's arbitrary movements and their goal choices. My findings in Chapter 4, further showed that 11-month-olds sensitivity to the shared movement repertoires of agents belonging to the same minimal group with each other modulated their expectation of efficient approach in goal-directed action pursuits.

In my empirical work for Chapter 4, critically, familiarization actions were performed in the presence of the test agent who presumably was interpreted as actively observing her social group partners (in Experiments 2 and 3). Given that imitative interactions between agents can inform infants' social evaluations, it is possible that infants expect imitative interactions to prevail between agents they can represent as belonging to the same social group. Likewise, they would not expect an outgroup party to engage in this imitative interaction. In a scenario where the test agent would not be present in the scene observing how her group members behave, perhaps there would be no modulatory effect of shared group movement repertoire on infants' expectation of efficiency. Yet if I am right in my hypothesis that sub-efficient actions carry normative weight elucidating the socially expected way the actions should be performed in one's social group, then future research should be able to replicate my findings even when the test agent was not there at the scene during the familiarization. This would indicate that infants do not only recognize imitative patterns between social group members' movement repertoires but that they expect a novel social group member to *know* how to act in the same context despite not viewing how the members of her social group acted before. Yet so far, the evidence we have only points out that infants do recognize imitative movement patterns between social agents (Powell & Spelke, 2018a, 2018b), could expect social agents to copy the arbitrary movements of their group members (Powell & Spelke, 2013), and that this expectation of shared group movement repertoire extends over object-directed actions interacting with infants' robust expectation of efficiency at the age of 11-months as Chapter 4 documented.

It is a possibility that preverbal infants' sensitivity to being imitated and also their social preferences and evaluations for those agents who imitatively interact with each other could scale up later in ontogeny serving two functions: it provides an epistemic benefit due to providing an opportunity to learn socially shared new action-means despite their causal opacity, and further it provides a social affiliation benefit due to the mean action acting as a social signal to the group members of the performer indicating (and strengthening the perception of) the performer's social affiliation and belongingness to her social group. I argue that infants' sensitivity to the imitative actions between agents belonging to the same social group even sometimes at the expense of efficacy serves as a foundational block for their developing understanding about the normative structure of human cultural activities.

There is indeed now evidence that as early as 18-months young children can entertain that it is not only shared experiences between themselves and their social partners that guides their interpretations of others' actions, but culturally shared and prescribed activities at group level (Schmidt, Rakoczy, & Tomasello, 2019). Upon learning novel and arbitrary actions from an ostensive adult informant, infants were presented with a third-party (a puppet) who performed a distinctively different action than the adult informant previously showed to the infants. Participants were found to communicatively intervene against the third party and undo its actions when it performed a different action in the condition when infants were told that the action ought to be performed in one particular way. In comparison, infants' communicative and behavioral interventions were rather low in the other condition when the experimenter called the infants' attention that there could be also other ways that the action could be performed but she just preferred to do it that way. Findings of Schmidt and his colleagues (2019) reveal that before their second year, infants do understand that there might be a *collective way* of doing something. They do not only form descriptive expectations about how a novel agent will act like but they generate normative expectations about how a novel agent should act like. Other studies also show that as they get older children show more robust protest behavior against third parties if they perform different actions than normatively prescribed (Rakoczy, Warneken, & Tomasello, 2008). Their protest behavior is also selective, they enforce norms on ingroup members but not on outgroup members (Schmidt, Rakoczy, & Tomasello, 2012), documenting that pre-school aged children generalize norms to social category members (also see Kalish, 2012; Kalish & Lawson, 2008).

As previous research documented, already before their second birthday infants understand that there is a collective reason behind the intentional action of an individual agent. I argued that being epistemically oriented to the informants speaking the same language as them contributes to young children's understanding of the collective intentionality behind the causally opaque actions in pedagogical transmission contexts, further providing evidence for the naïve learner about the socially constituted reason behind the action performance. However, I believe the precursors of this understanding could be found even earlier as documented by studies with pre-verbal infants. Pre-verbal infants' sensitivity to shared action movement repertoires of agents acting in a group context as I documented, along with the previous research revealing their sensitivity to imitative interactions between third parties (Powell & Spelke, 2018a, 2018b), suggest that these early expectations can be building blocks for understanding the conventional practices of cultural groups.

6.7 Concluding remarks

A crucial task of children is to acquire languages, beliefs, skills and practices of the cultural groups they live in (Csibra & Gergely, 2011; Tomasello, 2009; Disendruck & Markson, 2011). Critically, most forms of cultural knowledge, has no underlying physical causal mechanism enabling the observer to reflect on the behaviors, gestures, and words of others. For example, why the word *book* means what it means cannot be derived from perceptual input alone. Rather what gives this its referential meaning is the conventional agreement behind it (Diesendruck, 2005; principle of conventionality, Clark, 1988, 1990). Likewise, certain practices of cultural groups often do not have any observable cause and effect structure enabling naïve observers to make sense of the actions they observe in relation to their intended outcomes (Gergely & Csibra, 2006), such as in the case of conventional and ritualistic actions. Yet still the way we act in the world is also constrained by social conventional practices of our cultural communities that prescribe how we should do things in certain contexts. In some sense, the critical components of some of the actions we perform is not only concerned with what we do but also how we do it. Conventional practices we engage in different cultural groups are full of examples demonstrating the importance of the stylistic manners in the performance of seemingly instrumental actions with no perceivable adaptive value.

Socially stipulated conventional practices are a universal feature of human cultures (Legare & Nielsen, 2005; Legare & Harris, 2016). From early on, infants could rely on several cognitive biases that are likely to work in tandem with each other supporting their efficient learning of such cognitively opaque knowledge forms (Csibra & Gergely, 2006, 2007; Gergely & Csibra, 2005, 2006). The research presented in this dissertation focused on the role of the same-language cue in its epistemic benefit in supporting infants' and children's learning and adherence to cognitively opaque mean actions given that shared-language can signal the informants' representative cultural expertise. It further portrayed pre-verbal infants' sensitivity to the shared action repertoires between perceptually salient group of characters and documented how infants' early social categorization could inform their predictions about the conventionality of an action repertoire. Having a propensity to track shared movement repertoires of agents belonging to the same group, and later being able to differentially interpret the pedagogical demonstrations of cognitively opaque mean actions when they are delivered by linguistic ingroup members provides an insight into how young children become competent cultural members of their own communities.

Appendices

Appendix 1

Chapter 2

The number of head- and hand-touch actions performed by infants in the three conditions

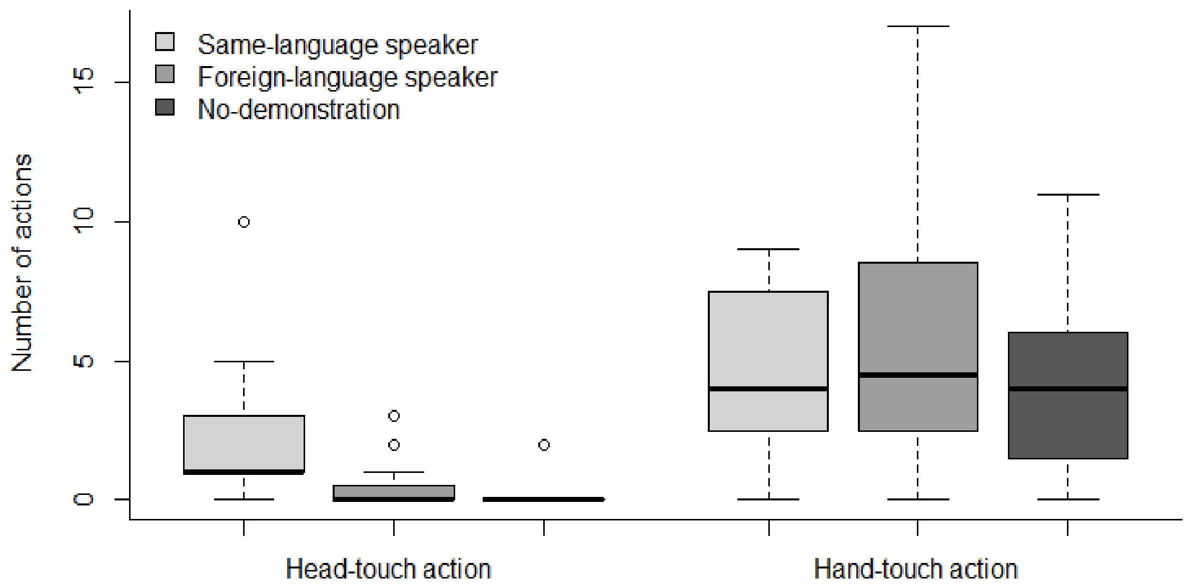

Figure 1. Boxplot for the number of head- and hand-touch actions performed by infants in the same-language speaker and foreign-language speaker conditions

As could be seen in Figure 1, the mean number of hand-touches infants performed was 4.56 (*SD* = 2.68) in the same-language speaker condition, 5.62 (*SD* = 4.35) in the foreign-language speaker condition, and it was 4.06 (*SD* = 3.27). in the no-demonstration condition. A Kruskal Wallis test revealed no difference in the frequency of hand-touches infants performed in three conditions (χ^2 (2) = .98, $p > .05$).

Mean number of head-touches infants performed was 2.12 (*SD* = 2.58) in the same-language speaker condition, .56 (*SD* = 1.09) in the foreign-language speaker condition, and it was .25 (*SD* = .68). in the no-demonstration condition. A Kruskal Wallis test revealed that the number of head-touches infants performed was statistically significantly different in three conditions (χ^2 (2) = 15.09, $p < .001$). We followed up this difference with independent samples non-parametric tests which indicated that the infants in the same language condition performed more head touch actions compared to the infants in the foreign language condition ($U = 59.5$, $p = .005$) and the infants in the no demonstration condition ($U = 43$, $p < .001$). There were no differences in head-touch action frequencies between the infants in the foreign language and no demonstration conditions ($U = 111$, $p = .36$).

However, note that there were more infants performing a head-touch action in the same-language speaker condition. Hence, we performed a post-hoc comparison that only compared the frequency of head-touch actions by the head-touch imitators in the three conditions (*n*'s, respectively: 13, 4, 2). Mean number of head touch actions by the head touch imitators in the same-language speaker condition was 2.61 ($SD = 2.63$), it was 2.25 ($SD = .96$) for the head touch imitators in the foreign language speaker condition and 2 ($SD = .0$) for the head touch imitators in the no demonstration condition. A Kruskal Wallis test indicated no difference in the frequency of head actions head-touch imitators performed in three conditions ($\chi^2(2) = .36$, $p > .87$).

Appendix 2

Chapter 4

Inter-rater reliability for exclusions

Experiment 1a. Individual efficiency:

Given the high exclusion rate mostly due to premature termination of test trials, a third coder, blinded to the type of the test trials, coded the video recordings of test trials from all participating infants ($N = 49$) to determine whether they met the inclusion criteria (please see the procedure in the main text). Percentage of agreement between the first two coders and the naïve coder was very high (97.92, Kappa = .96). Fifty six out of all 216 test trials (26%) were initially coded as terminating prematurely by the two experimenters during the experiment. The third coder blind to the test trials agreed with that decision 96% of the time. Disagreement about two trials was resolved through discussion.

Experiment 1b. Efficient action trajectories:

A third coder, blind to the condition, coded the video recordings of all trials from all tested infants ($N = 46$, except of six who were excluded due to the calibration error or the script crashing) to determine whether they met the inclusion criteria. Percentage of agreement between the first two coders and the naïve coder was high (97.3, Kappa = .95). Fifty six out of all 230 test trials (24.35%) were initially coded as terminating prematurely by the two experimenters. The third coder blind to the condition agreed with that decision 98% of the time. Disagreement about one trial was resolved through discussion.

Experiment 2. Ingroup inefficiency:

A third coder, blinded to the type of the test trial, coded the video recordings of all trials from all tested infants ($N = 52$) to determine whether they met the inclusion criteria. Percentage of agreement between the first two coders and the naïve coder was high (94.25, Kappa = .88).

Fifty one out of all 226 test trials (22.57%) were initially coded as terminating prematurely by the two experimenters. The third coder blind to the condition agreed with that decision 98% of the time. Disagreement about one trial was resolved through discussion.

Experiment 3. Ingroup efficiency:

A third coder, blind to the condition, coded the video recordings of all trials from all tested infants (N = 43) to determine whether they met the inclusion criteria. Percentage of agreement between the first two coders and the naïve coder was exceptionally high (97.13, Kappa = .94). Thirty one out of all 209 test trials (14.83%) were initially coded as terminating prematurely by the two experimenters. The third coder blind to the condition agreed with that decision 93% of the time. Disagreement about two trials was resolved through discussion.

Experiment 4. Outgroup efficiency:

A third coder, blind to the condition, coded the video recordings of all trials from all tested infants (N = 43) to determine whether they met the inclusion criteria. Percentage of agreement between the first two coders and the naïve coder was exceptionally high (98.43%, Kappa = .97). Sixty nine out of all 191 test trials (36.13%) were initially coded as terminating prematurely by the two experimenters. The third coder blind to the condition agreed with that decision 97.1% of the time. Disagreement about two trials was resolved through discussion.

Looking behavior at the familiarization events across experiments

Below we list looking time durations of infants to familiarization events in each experiment.

Experiment 1a. Individual efficiency:

Table 1. Mean looking times to familiarization events with the standard deviations in parentheses, and the duration of the movies used in these familiarization phases along with the

percentage of time the infants looked at the familiarization events in Experiment 1, *individual efficiency*. Note the sample size is different for each familiarization event.

	Mean (*SD*)	Duration	% viewed	*N*
Group induction	50.55 (*6.73*)	57.4	88.06	24
Action familiarization	48.59 (*8.44*)	62.3	78	24
Group reminder & second action familiarization	38.7 (*11.27*)	54.9	70.49	23
Second Group reminder & third action familiarization	38.12 (*11.3*)	54.9	69.43	15

Experiment 1b. Efficient action trajectories:

Table 2. Mean looking times to familiarization events with the standard deviations in parentheses, and the duration of the movies used in these familiarization phases along with the percentage of time the infants looked at the familiarization events in Experiment 4, *efficient action trajectories*. Note the sample size is different for each familiarization event.

	Mean (*SD*)	Duration	% viewed	*N*
Action familiarization	55.12 (5.*91*)	62.3	96.03	24
Second action familiarization	25.03 (4.*77*)	31.5	79.46	22
Third action familiarization	22.66 (6.*71*)	31.5	71.93	18

Experiment 2. Ingroup inefficiency:

Table 3. Mean looking times to familiarization events with the standard deviations in parentheses, and the duration of the movies used in these familiarization phases along with the percentage of time the infants looked at the familiarization events in Experiment 2, *ingroup inefficiency*. Note the sample size is different for each familiarization event.

	Mean (*SD*)	Duration	% viewed	*N*
Group induction	49.67 (*6.98*)	57.4	86.53	24
Action familiarization	51.45 (9.*38*)	62.3	82.58	24
Group reminder & second action familiarization	40.73 (7.*8*)	54.9	74.19	24
Second Group reminder & third action familiarization	36.16 (*12.02*)	54.9	65.85	16

Experiment 3. Ingroup efficiency:

Table 4. Mean looking times to familiarization events with the standard deviations in parentheses, and the duration of the movies used in these familiarization phases along with the

percentage of time the infants looked at the familiarization events in Experiment 3, *ingroup efficiency*. Note the sample size is different for each familiarization event.

	Mean (*SD*)	Duration	% viewed	*N*
Group induction	51.41 (*6.38*)	57.4	89.56	24
Action familiarization	53.09 (6.*83*)	62.3	85.21	24
Group reminder & second action familiarization	43.98 (7.*1*)	54.9	80.1	21
Second Group reminder & third action familiarization	40.92 (8.*7*)	54.9	74.53	17

Experiment 4. Outgroup efficiency:

Table 5. Mean looking times to familiarization events with the standard deviations in parentheses, and the duration of the movies used in these familiarization phases along with the percentage of time the infants looked at the familiarization events in Experiment 5, *outgroup efficient*. Note the sample size is different for each familiarization event.

	Mean (*SD*)	Duration	% viewed	*N*
Group induction	47.67 (10.*61*)	57.4	83.05	24
Action familiarization	54.54 (6.*96*)	62.3	87.54	24
Group reminder & second action familiarization	38.62 (13.06)	54.9	70.35	24
Second Group reminder & third action familiarization	40.89 (*13.48*)	54.9	74.48	15

Appendix 3

Chapter 5

English translations of the model's utterances during the pretend phone-conversation were as following (text 1 was used in experiments 1 and 3, and text 1 and text 2 were used in experiment 2):

Text 1: Yes mom, I am listening. Oh you cannot find your keys? Have you tried checking the cupboard at the entrance? Yes go on, check the first drawer. Alright have you got them now? Ok no problem then. Enjoy your shopping, kisses. See you later in the evening.

Text 2: Yes yes, I have heard about the birthday party. I am not sure if I can make it this weekend though. It was Saturday, right? Are you going? I have to visit my parents in the countryside. I will let you know.

Imitation of causally relevant action

We also coded children's imitation of the demonstrated causally relevant action in three experiments. For each experiment, data from five children were further coded by a second coder who was naïve to the condition and the hypotheses of the study. Inter-coder reliability was excellent for both 5-year-olds (Cohen's Kappa = 1, $p < .001$) and 6-year-olds (Cohen's Kappa = 1, $p < .001$) in Experiment 1. Inter-coder reliability was also perfect for Experiments 2 and 3 (Cohen's Kappa for both experiments = 1, ps < .001).

Experiment 1. Results: Causally relevant action imitation

All children imitated the causally relevant action in both imitation rounds, except one 5-year-old in both imitation rounds of the *same + foreign* condition, and another 5-year old in the first imitation round of the *foreign + same condition*. These two children retrieved the sticker by pulling the platform out with their hands.

Experiment 2. Results: Causally relevant action imitation

Twelve children in the first imitation round and 14 in the second round imitated the demonstrated causally relevant action. The remaining children retrieved the sticker by using their hands in pulling the platform out of the sticker dispenser.

Experiment 3. Results: Causally relevant action imitation

All children except one in each condition imitated the demonstrated causally relevant action in the first imitation round. All but one in the *same + foreign* condition imitated the causally relevant action in the second round. Those three children used their hands in pulling the platform out of the sticker dispenser in retrieving the sticker.

References

Agnetta, B., & Rochat, P. (2004). Imitative games by 9-, 14-, and 18-month-old infants. *Infancy*, *6*(1), 1-36.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.

Bahrick, L. E., Netto, D., & Hernandez-Keif, M. (1998). Intermodal perception of adult and child faces and voices by infants. *Child Development*, *69*(5), 1263-1275.

Baker, M. C. (2001). *The atoms of language: The mind's hidden rules of grammar*. New York, NY: Basic Books.

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4), 1-10.

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*(3), 329-349.

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. *Psychological Science*, *17*(2), 159-163.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497-529.

Begus, K., Curioni, A., Gergely, G., & Knoblich, G. (2018, January). Joint goal representation in infants: a fNIRS study. Poster presented at Budapest CEU Conference on Cognitive Development, Budapest, Hungary.

Begus, K., Gliga, T., & Southgate, V. (2016). Infants' preferences for native speakers are associated with an expectation of information. *Proceedings of the National Academy of Sciences*, *113*(44), 12397-12402.

Bekkering, H., Wohlschlager, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *The Quarterly Journal of Experimental Psychology: Section A*, *53*(1), 153-164.

Bennett, M., Barrett, M., Karakozov, R., Kipiani, G., Lyons, E., Pavlenko, V., & Riazanova, T. (2004). Young children's evaluations of the ingroup and of outgroups: A multi-national study. *Social Development*, *13*(1), 124-141.

Bian, L., Sloane, S., & Baillargeon, R. (2018). Infants expect ingroup support to override fairness when resources are limited. *Proceedings of the National Academy of Sciences*, *115*(11), 2705-2710.

Billig, M., & Tajfel, H. (1973). Social categorization and similarity in intergroup behaviour. *European Journal of Social Psychology*, *3*(1), 27-52.

Birch, S. A., Akmal, N., & Frampton, K. L. (2010). Two-year-olds are vigilant of others' non-verbal cues to credibility. *Developmental Science*, *13*(2), 363-369.

Birnbaum, D., Deeb, I., Segall, G., Ben-Eliyahu, A., & Diesendruck, G. (2010). The development of social essentialism: The case of Israeli children's inferences about Jews and Arabs. *Child Development*, *81*(3), 757-777.

Biro, S., & Leslie, A. M. (2007). Infants' perception of goal-directed actions: development through cue-based bootstrapping. *Developmental Science*, *10*(3), 379-398.

Biro, S., Verschoor, S., & Coenen, L. (2011). Evidence for a unitary goal concept in 12-month-old infants. *Developmental Science*, *14*(6), 1255-1260.

Blaustein, A. R. (1983). Kin recognition mechanisms: phenotypic matching or recognition alleles?. *The American Naturalist*, *121*(5), 749-754.

Bloch, M. (1974). Symbols, Song, Dance and Features of Articulation Is religion an extreme form of traditional authority?. *European Journal of Sociology/Archives Européennes de Sociologie*, *15*(1), 54-81.

Bloch, M. (2004). Bringing ritual to mind: Psychological foundations of cultural forms. *Journal of the Royal Anthropological Institute*, *10*(1), 202-204.

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, *120*(3), 322-330.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, *108*, 10918-10925.

Boyer, P., & Liénard, P. (2006). Why ritualized behavior? Precaution systems and action parsing in developmental, pathological and cultural rituals. *Behavioral and Brain Sciences*, *29*(6), 595-613.

Boyer, P., & Liénard, P. (2020). Ingredients of "rituals" and their cognitive underpinnings. *Philosophical Transactions of the Royal Society*. DOI: 10.1098/rstb.2019-0439

Bowern, C. (2010). Correlates of Language Change in Hunter-Gatherer and Other 'Small' Languages. *Language and Linguistics Compass*, *4*(8), 665-679.

Brandone, A. C., & Wellman, H. M. (2009). You can't always get what you want: Infants understand failed goal-directed actions. *Psychological Science*, *20*(1), 85-91.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate?. *Journal of Social Issues*, *55*(3), 429-444.

Brooker, I., & Poulin-Dubois, D. (2013). Is a bird an apple? The effect of speaker labeling accuracy on infants' word learning, imitation, and helping behaviors. *Infancy*, *18*, E46-E68.

Brown, R., & Pehrson, S. (2019). *Group processes: Dynamics within and between groups*. NJ: Wiley-Blackwell.

Brugger, A., Lariviere, L. A., Mumme, D. L., & Bushnell, E. W. (2007). Doing the right thing: Infants' selection of actions to imitate from observed event sequences. *Child Development*, *78*(3), 806-824.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., ... & Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, *13*(2), 400-404.

Buccino, G., Binkofski, F., & Riggio, L. (2004). The mirror neuron system and action recognition. *Brain and Language*, *89*(2), 370-376.

Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition*, *120*(3), 331-340.

Bulbulia, J. (2004). The cognitive and evolutionary psychology of religion. *Biology and Philosophy, 9(5),* 655-686.

Burdett, E. R., McGuigan, N., Harrison, R., & Whiten, A. (2018). The interaction of social and perceivable causal factors in shaping 'over-imitation'. *Cognitive Development*, *47*, 8-18.

Bussey, K., & Bandura, A. (1984). Influence of gender constancy and social power on sex-linked modeling. *Journal of Personality and Social Psychology*, *47*(6), 1292.

Butler, L. P., & Markman, E. M. (2012). Preschoolers use intentional and pedagogical cues to guide inductive inferences and exploration. *Child Development*, *83*(4), 1416-1428.

Butler, L. P., Schmidt, M. F., Bürgel, J., & Tomasello, M. (2015). Young children use pedagogical cues to modulate the strength of normative inferences. *British Journal of Developmental Psychology*, *33*(4), 476-488.

Buttelmann, D., & Böhm, R. (2014). The ontogeny of the motivation that underlies in-group bias. *Psychological Science*, *25*(4), 921-927.

Buttelmann, D., Carpenter, M., Call, J., & Tomasello, M. (2008). Rational tool use and tool choice in human infants and great apes. *Child Development*, *79*(3), 609-626.

Buttelmann, D., Schieler, A., Wetzel, N., & Widmann, A. (2017). Infants' and adults' looking behavior does not indicate perceptual distraction for constrained modelled actions− An eye-tracking study. *Infant Behavior and Development*, *47*, 103-111.

Buttelmann, D., & Zmyj, N. (2012). Evaluating the empirical evidence for the two-stage-model of infant imitation. A commentary on Paulus, Hunnius, Vissers, and Bekkering (2011). *Frontiers in Psychology*, *3*, 512.

Buttelmann, D., Zmyj, N., Daum, M., & Carpenter, M. (2013). Selective imitation of in-group over out-group members in 14-month-old infants. *Child Development*, *84*(2), 422-428.

Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences*, *21*(5), 667-684.

Call, J., & Carpenter, M. (2002). Three sources of information in social learning. In K. Dautenhahn & C. Nehaniv (Eds.), *Imitation in animals and artifacts* (pp. 211-228). Cambridge, MA: MIT Press.

Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science*, *3*(1), 14-25.

Campbell, M. W., & De Waal, F. B. (2011). Ingroup-outgroup bias in contagious yawning by chimpanzees supports link to empathy. *PloS One*, *6*(4).

Carey, S. (2009). *The origin of concepts*. Oxford, UK: Oxford University Press.

Carpenter, M. (2006). Instrumental, social, and shared goals and intentions in imitation. In S.J. Rogers & J. Williams (Eds.), *Imitation and the development of the social mind: Lessons from typical development and autism* (pp. 48–70). New York: Guilford.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen-through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, *21*(2), 315-330.

Carpenter, M., & Call, J. (2009). Comparing the imitative skills of children and nonhuman apes. *Revue de Primatologie, 1.*

Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve-and 18-month-olds copy actions in terms of goals. *Developmental Science*, *8*(1), F13-F20.

Carpenter, M., Uebel, J., & Tomasello, M. (2013). Being mimicked increases prosocial behavior in 18-month-old infants. *Child Development*, *84*(5), 1511-1518.

Chalik, L., Rivera, C., & Rhodes, M. (2014). Children's use of categories and mental states to predict social behavior. *Developmental Psychology*, *50*(10), 2360.

Chen, M. L., Kiraly, I. & Gergely, G. (2012, January). *Relevance-guided imitation and emulation in 14-month-olds.* Paper presented at Budapest CEU Conference on Cognitive Development, Budapest, Hungary.

Chen, M. L., & Waxman, S. R. (2013). "Shall we blick?": Novel words highlight actors' underlying intentions for 14-month-old infants. *Developmental Psychology*, *49*(3), 426.

Cimpian, A., & Scott, R. M. (2012). Children expect generic knowledge to be widely shared. *Cognition*, *123*(3), 419-433.

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

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

Clegg, J. M., & Legare, C. H. (2016). Instrumental and conventional interpretations of behavior are associated with distinct outcomes in early childhood. *Child Development*, *87*(2), 527-542.

Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences*, *7*(4), 173-179.

Cruz-Khalili, A., Bettencourt, K., Kohn, C. S., Normand, M. P., & Schlinger, H. D. (2019). Use of Repeated Within-Subject Measures to Assess Infants' Preference for Similar Others. *Frontiers in Psychology*, *10*.

Csibra, G. (2003). Teleological and referential understanding of action in infancy. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 447-458.

Csibra, G. (2008a). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition*, *107*(2), 705-717.

Csibra, G. (2008b). Action mirroring and action understanding: An alternative account. In P. Haggard, Y. Rosetti, & M. Kawato (Eds.), *Sensorymotor foundations of higher cognition. Attention and performance XXII* (pp. 435–459). Oxford: Oxford University Press.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, *25*(2), 141-168.

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*(1), 111-133.

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: the case for pedagogy. In Y. Munakata & Johnson, M.H. (Eds.), *Processes of change in brain and cognitive development (Attention and performance)* (Vol. 21, pp. 249-274). Oxford: Oxford University Press.

Csibra, G., & Gergely, G. (2007). 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, *124*(1), 60-78.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148-153.

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1149-1157.

Csibra, G., Gergely, G., Bíró, S., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: the perception of 'pure reason' in infancy. *Cognition*, *72*(3), 237-267.

Csibra, G., Hernik, M., Mascaro, O., Tatone, D., & Lengyel, M. (2016). Statistical treatment of looking-time data. *Developmental Psychology*, *52*(4), 521.

Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences*, *15*(5), 218-226.

Cohen, E. (2012). The evolution of tag-based cooperation in humans: The case for accent. *Current Anthropology, 53*(5), 588-616.

Daum, M. M., Prinz, W., & Aschersleben, G. (2008). Encoding the goal of an object-directed but uncompleted reaching action in 6-and 9-month-old infants. *Developmental Science*, *11*(4), 607-619.

Davies, A. F. (1968). The child's discovery of nationality. *The Australian and New Zealand Journal of Sociology*, *4*(2), 107-125.

DeJesus, J. M., Hwang, H. G., Dautel, J. B., & Kinzler, K. D. (2018). "American= English Speaker" Before "American= White": The Development of Children's Reasoning About Nationality. *Child Development*, *89*(5), 1752-1767.

Diesendruck, G. (2005). The principles of conventionality and contrast in word learning: An empirical examination. *Developmental Psychology*, *41*(3), 451.

Diesendruck, G., Carmel, N., & Markson, L. (2010). Children's sensitivity to the conventionality of sources. *Child Development*, *81*(2), 652-668.

Diesendruck, G., Goldfein-Elbaz, R., Rhodes, M., Gelman, S., & Neumark, N. (2013). Cross-cultural differences in children's beliefs about the objectivity of social categories. *Child Development*, *84*(6), 1906-1917.

Diesendruck, G., & HaLevi, H. (2006). The role of language, appearance, and culture in children's social category-based induction. *Child Development*, *77*(3), 539-553.

Diesendruck, G., & Markson, L. (2011). Children's assumption of the conventionality of culture. *Child Development Perspectives*, *5*(3), 189-195.

DiYanni, C. J., Corriveau, K. H., Kurkul, K., Nasrini, J., & Nini, D. (2015). The role of consensus and culture in children's imitation of inefficient actions. *Journal of Experimental Child Psychology*, *137*, 99-110.

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliation in children. *Child Development*, *82*(3), 793-811.

Egyed, K., Király, I., & Gergely, G. (2013). Communicating shared knowledge in infancy. *Psychological Science*, *24*(7), 1348-1353.

Elsner, B., & Aschersleben, G. (2003). Do I get what you get? Learning about the effects of self-performed and observed actions in infancy. *Consciousness and Cognition*, *12*(4), 732-751.

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 229.

Elsner, B., Pfeifer, C., Parker, C., & Hauf, P. (2013). Infants' perception of actions and situational constraints: An eye-tracking study. *Journal of Experimental Child Psychology*, *116*(2), 428-442.

Evans, C. L., Laland, K. N., Carpenter, M., & Kendal, R. L. (2018). Selective copying of the majority suggests children are broadly "optimal-" rather than "over-" imitators. *Developmental Science*, *21*(5), e12637.

Falk, C. F., Heine, S. J., & Takemura, K. (2014). Cultural variation in the minimal group effect. *Journal of Cross-Cultural Psychology*, *45*(2), 265-281.

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, *454*(7208), 1079-1083.

Frank, R. H. 1988. *Passions Within Reason: The strategic Role of the Emotions.* New York: W.W. Norton.

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

Gattis, M. (2002). Imitation is mediated by many goals, not just one. *Developmental Science, 5*, 27–29.

Gattis, M., Bekkering, H., & Wohlschläger, A. (2002). Goal-directed imitation. In A. N. Meltzoff & W. Prinz (Eds.), *The imitative mind: Development, evolution, and brain bases* (pp. 183–205). Cambridge: Cambridge University Press.

Gazzola, V., & Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cerebral Cortex*, *19*(6), 1239-1255.

Gellén, K., & Buttelmann, D. (2017). Fourteen-month-olds adapt their imitative behavior in light of a model's constraints. *Child Development Research*.

Gellén, K., & Buttelmann, D. (2019). Rational imitation declines within the second year of life: Changes in the function of imitation. *Journal of Experimental Child Psychology*, *185*, 148- 163.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York: Oxford University Press.

Gergely, G. (2013). Ostensive communication and cultural learning: The natural pedagogy hypothesis. In J. Metcalfe & H. S. Terrace (Eds.), *Agency and Joint Attention* (pp. 139-151). Oxford: Oxford University Press.

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*(6873), 755-755.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287-292.

Gergely, G., & Csibra, G. (2005). The social construction of the cultural mind: Imitative learning as a mechanism of human pedagogy. *Interaction Studies*, *6*(3), 463-481.

Gergely, G., & Csibra, G. (2006). Sylvia's recipe: The role of imitation and pedagogy in the transmission of cultural knowledge. In N.J. Enfield & S.C. Levinson (Eds.), *Roots of human sociality: Culture, cognition, and human interaction* (pp. 229-255). Oxford: Berg Press.

Gergely, G., Egyed, K., & Király, I. (2007). On pedagogy. *Developmental Science, 10*, 139–146.

Gergely, G., & Jacob, P. (2012). Reasoning about instrumental and communicative agency in human infancy. In J. B. Benson, F. Xu & T. Kushnir (Eds.), *Advances in child development and behavior. Vol 43: Rational constructivism in cognitive development* (pp. 59-94). Waltham, MA: Academic Press/Elsevier.

Gergely, G., & Jacob, P. (2013). Reasoning about instrumental and communicative agency in human infancy. In J. B. Benson (Ser. Ed.), F. Xu, & T. Kushnir (Vol. Eds.), *Rational constructivism in cognitive development* (pp. 59–94). San Diego: Elsevier.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165-193.

Gil-White, F.J. (2001). Are ethnic groups biological 'species' to the human brain? Essentialism in our cognition of some social categories. *Current Anthropology, 42*, 515–554.

Giles, H., & Billings, A. C. (2004). Assessing language attitudes: Speaker evaluation studies. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 187-209). Malden, MA: Blackwell.

Gleissner, B., Meltzoff, A. N., & Bekkering, H. (2000). Children's coding of human action: cognitive factors influencing imitation in 3-year-old. *Developmental Science*, *3*(4), 405-414.

Gramzow, R. H., & Gaertner, L. (2005). Self-esteem and favoritism toward novel in-groups: the self as an evaluative base. *Journal of Personality and Social Psychology*, *88*(5), 801.

Gruber, T., Deschenaux, A., Frick, A., & Clément, F. (2019). Group membership influences more social identification than social learning or overimitation in children. *Child Development*, *90*(3), 728-745.

Guderian, S., Schott, B. H., Richardson-Klavehn, A., & Düzel, E. (2009). Medial temporal theta state before an event predicts episodic encoding success in humans. *Proceedings of the National Academy of Sciences*, *106*(13), 5365-5370.

Hamilton, A. F. (2009). Research review: Goals, intentions and mental states: Challenges for theories of autism. *Journal of Child Psychology and Psychiatry*, *50*(8), 881-892.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. *Journal of Theoretical Biology*, *7*(1), 17-52.

Hansen, K., Rakić, T., & Steffens, M. C. (2018). Foreign-looking native-accented people: More competent when first seen rather than heard. *Social Psychological and Personality Science*, *9*(8), 1001-1009.

Hauf, P., Aschersleben, G., & Prinz, W. (2007). Baby do–baby see!: How action production influences action perception in infants. *Cognitive Development*, *22*(1), 16-32.

Haun, D. B., & Call, J. (2008). Imitation recognition in great apes. *Current Biology*, *18*(7), R288-R290.

Haun, D. B. M., & Over, H. (2015). A homophily-based account of human culture. In P.J. Richerson & M. H. Christiansen (Eds.), *Cultural evolution: Society, technology, language, and religion* (pp. 117–130). Cambridge, MA: MIT Press

Henderson, A. M., Wang, Y., Matz, L. E., & Woodward, A. L. (2013). Active experience shapes 10-month-old infants' understanding of collaborative goals. *Infancy*, *18*(1), 10-39.

Henderson, A. M., & Woodward, A. L. (2011). "Let's work together": What do infants understand about collaborative goals?. *Cognition*, *121*(1), 12-21.

Henrich, J. (2009). The evolution of costly displays, cooperation and religion: Credibility enhancing displays and their implications for cultural evolution. *Evolution and Human Behavior*, *30*(4), 244-260.

Henrich, N., & Henrich, J. P. (2007). *Why humans cooperate: A cultural and evolutionary explanation*. Oxford: Oxford University Press.

Hepach, R., & Westermann, G. (2016). Pupillometry in infancy research. *Journal of Cognition and Development*, *17*(3), 359-377.

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

Hernik, M., & Gergely, G. (2015). To what adaptive problems is human teaching a solution? *Behavioral and Brain Sciences*, *38*.

Herrmann, P. A., Legare, C. H., Harris, P. L., & Whitehouse, H. (2013). Stick to the script: The effect of witnessing multiple actors on children's imitation. *Cognition*, *129*(3), 536-543.

Heyman, G. D., & Gelman, S. A. (2000). Beliefs about the origins of human psychological traits. *Developmental Psychology, 36*(5), 663–678.

Hirschfeld, L.A. (1996). *Race in the making: Cognition, culture, and the child's construction of human kinds*. Cambridge, MA: The MIT Press.

Hirschfeld, L.A., & Gelman, S.A. (1997). What young children think about the relationship between language variation and social difference. *Cognitive Development*, *12*, 213–238.

Hobson, R. P., & Hobson, J. A. (2008). Dissociable aspects of imitation: A study in autism. *Journal of Experimental Child Psychology*, *101*(3), 170-185.

Hobson, R. P., & Lee, A. (1999). Imitation and identification in autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *40*(4), 649-659.

Hoehl, S., Keupp, S., Schleihauf, H., McGuigan, N., Buttelmann, D., & Whiten, A. (2019). 'Over-imitation': A review and appraisal of a decade of research. *Developmental Review*, *51*, 90-108.

Hoehl, S., Zettersten, M., Schleihauf, H., Grätz, S., & Pauen, S. (2014). The role of social interaction and pedagogical cues for eliciting and reducing overimitation in preschoolers. *Journal of Experimental Child Psychology*, *122*, 122-133.

Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (Pan troglodytes) and children (Homo sapiens). *Animal Cognition*, *8*(3), 164-181

Howard, L. H., Carrazza, C., & Woodward, A. L. (2014). Neighborhood linguistic diversity predicts infants' social learning. *Cognition*, *133*(2), 474-479.

Howard, L. H., Henderson, A. M., Carrazza, C., & Woodward, A. L. (2015). Infants' and young children's imitation of linguistic in-group and out-group informants. *Child Development*, *86*(1), 259-275.

Humphrey, C., & Laidlaw, J. (1994). *The archetypal action of ritual: A theory of ritual illustrated by the Jain rite of worship*. Oxford: Clarendon Press.

Irons, W. G. (2001). Religion as a hard-to-fake sign of commitment. In R.M. Nesse (Ed.), *Evolution and the capacity for commitment* (pp. 292–309). New York: Russell Sage Foundation.

Jara-Ettinger, J., Floyd, S., Tenenbaum, J. B., & Schulz, L. E. (2017). Children understand that agents maximize expected utilities. *Journal of Experimental Psychology: General*, *146*(11), 1574.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, *20*(8), 589-604.

Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition*, *140*, 14-23.

Jin, K. S., & Baillargeon, R. (2017). Infants possess an abstract expectation of ingroup support. *Proceedings of the National Academy of Sciences*, *114*(31), 8199-8204.

Jin, K.S., & Baillargeon, R. Manuscript in preparation. Infants' expectations about harm avoidance within and between moral circles.

Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science*.

Kalish, C. W. (2012). Generalizing norms and preferences within social categories and individuals. *Developmental Psychology*, *48*(4), 1133.

Kalish, C. W., & Lawson, C. A. (2008). Development of social category representations: Early appreciation of roles and deontic relations. *Child Development*, *79*(3), 577-593.

Kamewari, K., Kato, M., Kanda, T., Ishiguro, H., & Hiraki, K. (2005). Six-and-a-half-month-old children positively attribute goals to human action and to humanoid-robot motion. *Cognitive Development*, *20*(2), 303-320.

Kapitány, R., & Nielsen, M. (2015). Adopting the ritual stance: The role of opacity and context in ritual and everyday actions. *Cognition*, *145*, 13-29.

Kapitány, R., & Nielsen, M. (2017). The ritual stance and the precaution system: The role of goal-demotion and opacity in ritual and everyday actions. *Religion, Brain & Behavior*, *7*(1), 27-42.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science*, *18*(12), 1084-1089.

Kenward, B. (2012). Over-imitating preschoolers believe unnecessary actions are normative and enforce their performance by a third party. *Journal of Experimental Child Psychology*, *112*(2), 195-207.

Kenward, B., Karlsson, M., & Persson, J. (2011). Over-imitation is better explained by norm learning than by distorted causal learning. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1709), 1239-1246.

Keupp, S., Behne, T., & Rakoczy, H. (2013). Why do children overimitate? Normativity is crucial. *Journal of Experimental Child Psychology*, *116*(2), 392-406.

Keupp, S., Behne, T., Zachow, J., Kasbohm, A., & Rakoczy, H. (2015). Over-imitation is not automatic: Context sensitivity in children's overimitation and action interpretation of causally irrelevant actions. *Journal of Experimental Child Psychology*, *130*, 163-175.

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron*, *88*(3), 449-460.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS One*, *7*(5), e36399.

Kinzler, K. D., Corriveau, K. H., & Harris, P. L. (2011). Preschoolers' use of accent when deciding which informant to trust. *Developmental Science*, *14*, 106-111.

Kinzler, K. D., & Dautel, J. B. (2012). Children's essentialist reasoning about language and race. *Developmental Science*, *15*(1), 131-138.

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, *104*(30), 12577-12580.

Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition*, *27*(4), 623-634.

Kinzler, K. D., & Liberman, Z. (2017). Infants' inferences about language are social. *Proceedings of the National Academy of Sciences*, *114*(19), E3753-E3754.

Kinzler, K. D., & Spelke, E. S. (2011). Do infants show social preferences for people differing in race?. *Cognition*, *119*(1), 1-9.

Király, I. (2009a). The effect of the model's presence and of negative evidence on infants' selective imitation. *Journal of Experimental Child Psychology*, *102*(1), 14-25.

Király, I. (2009b). Memories for events in infants: goal relevant action coding. In T. Striano & V. Reid (Eds.), *Social cognition: Development, Neuroscience and Autism* (pp. 113–128). London: Wiley-Blackwell.

Király, I., Csibra, G., & Gergely, G. (2013). Beyond rational imitation: Learning arbitrary means actions from communicative demonstrations. *Journal of Experimental Child Psychology*, *116*(2), 471-486.

Kovacs, A. M., Téglás, E., Gergely, G., & Csibra, G. (2017). Seeing behind the surface: communicative demonstration boosts category disambiguation in 12-month-olds. *Developmental Science*, *20*(6), e12485.

Krieger, A. A., Aschersleben, G., Sommerfeld, L., & Buttelmann, D. (2020). A model's natural group membership affects over-imitation in 6-year-olds. *Journal of Experimental Child Psychology*, *192*, 104783.

Krieger, A. A., Möller, C., Zmyj, N., & Aschersleben, G. (2016). Tom is not more likely to imitate Lisa than Ying: The influence of a model's race indicated by physical appearance on children's imitation. *Frontiers in Psychology*, *7*, 972.

Kupán, K., Király, I., Kupán, K., Krekó, K., Miklósi, Á., & Topál, J. (2017). Interacting effect of two social factors on 18-month-old infants' imitative behavior: Communicative cues and demonstrator presence. *Journal of Experimental Child Psychology*, *161*, 186-194.

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, *98*(26), 15387-15392.

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious?. *Perspectives on Psychological Science*, *7*(1), 18-27.

Legare, C. H. (2019). The Development of Cumulative Cultural Learning. *Annual Review of Developmental Psychology*, *1*, 119-147.

Legare, C. H., & Harris, P. L. (2016). The ontogeny of cultural learning. *Child Development*, *87*(3), 633-642.

Legare, C. H., & Herrmann, P. A. (2013). Cognitive consequences and constraints on reasoning about ritual. *Religion, Brain & Behavior*, *3*(1), 63-65.

Legare, C. H., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends in Cognitive Sciences*, *19*(11), 688-699.

Legare, C. H., & Souza, A. L. (2012). Evaluating ritual efficacy: Evidence from the supernatural. *Cognition*, *124*(1), 1-15.

Legare, C. H., & Souza, A. L. (2014). Searching for control: Priming randomness increases the evaluation of ritual efficacy. *Cognitive Science*, *38*(1), 152-161.

Legare, C. H., Watson-Jones, R. E. (2015). The evolution and ontogeny of ritual. In Buss, D. M. (Ed.), *The handbook of evolutionary psychology* (pp. 829–847). Hoboken, NJ: John Wiley & Sons.

Legare, C. H., & Wen, N. (2014). The effects of ritual on the development of social group cognition. *International Society for the Study of Behavioral Development*, *2*(66), 9-12

Legare, C. H., Wen, N. J., Herrmann, P. A., & Whitehouse, H. (2015). Imitative flexibility and the development of cultural learning. *Cognition*, *142*, 351-361.

Leinbach, M. D., & Fagot, B. I. (1993). Categorical habituation to male and female faces: Gender schematic processing in infancy. *Infant Behavior and Development*, *16*(3), 317-332.

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, *46*(6), 1093-1096.

Lev-Ari, S., & Keysar, B. (2012). Less-detailed representation of non-native language: Why non-native speakers' stories seem more vague. *Discourse Processes*, *49*(7), 523-538.

Li, Y., Liao, Y., Cheng, Y., & He, J. (2019). Group conquers efficacy: Preschoolers' imitation under conflict between minimal group membership and behavior efficacy. *PloS One*, *14*(9).

Lieberman, D., Oum, R., & Kurzban, R. (2008). The family of fundamental social categories includes kinship: Evidence from the memory confusion paradigm. *European Journal of Social Psychology*, *38*(6), 998-1012.

Liberman, Z., Kinzler, K. D., & Woodward, A. L. (2018). The early social significance of shared ritual actions. *Cognition*, *171*, 42-51.

Liberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). The origins of social categorization. *Trends in Cognitive Sciences*, *21*(7), 556-568.

Lickel, B., Hamilton, D. L., Wieczorkowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology*, *78*(2), 223.

Liénard, P., & Boyer, P. (2006). Whence collective rituals? A cultural selection model of ritualized behavior. *American Anthropologist*, *108*(4), 814-827.

Liu, S., Brooks, N. B., & Spelke, E. S. (2019). Origins of the concepts cause, cost, and goal in prereaching infants. *Proceedings of the National Academy of Sciences*, *116*(36), 17747-17752.

Liu, S., & Spelke, E. S. (2017). Six-month-old infants expect agents to minimize the cost of their actions. *Cognition*, *160*, 35-42.

Liu, S., Ullman, T., Tenenbaum, J., & Spelke, E. (2019). Dangerous ground: Thirteen-month-old infants are sensitive to peril in other people's actions. DOI: 10.31234/osf.io/rvydk

Luo, Y. (2011). Three-month-old infants attribute goals to a non-human agent. *Developmental Science*, *14*(2), 453-460.

Lyons, D. E., Damrosch, D. H., Lin, J. K., Macris, D. M., & Keil, F. C. (2011). The scope and limits of overimitation in the transmission of artefact culture. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *366*(1567), 1158-1167.

Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences*, *104*(50), 19751-19756.

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, *51*(1), 93-120.

Mahajan, N., & Wynn, K. (2012). Origins of "us" versus "them": Prelinguistic infants prefer similar others. *Cognition*, *124*(2), 227-233.

Marshall, P. J., & Meltzoff, A. N. (2014). Neural mirroring mechanisms and imitation in human infants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1644), 20130620.

Marno, H., Guellai, B., Vidal, Y., Franzoi, J., Nespor, M., & Mehler, J. (2016). Infants' selectively pay attention to the information they receive from a native speaker of their language. *Frontiers in Psychology*, *7*, 1150.

Marsh, L., Pearson, A., Ropar, D., & Hamilton, A. (2013). Children with autism do not overimitate. *Current Biology*, *23*(7), R266-R268.

Marsh, L. E., Ropar, D., & Hamilton, A. F. D. C. (2019). Are you watching me? The role of audience and object novelty in overimitation. *Journal of Experimental Child Psychology*, *180*, 123-130.

McGuigan, N., Makinson, J., & Whiten, A. (2011). From over-imitation to super-copying: Adults imitate causally irrelevant aspects of tool use with higher fidelity than young children. *British Journal of Psychology*, *102*(1), 1-18.

McLoughlin, N., Tipper, S. P., & Over, H. (2018). Young children perceive less humanness in outgroup faces. *Developmental Science*, *21*(2), e12539.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*(2), 143-178.

Meltzoff, A. N. (1988). Infant imitation and memory: Nine-month-olds in immediate and deferred tests. *Child Development, 59*, 217–225.

Meltzoff, A. N. (1990). Foundations for developing a concept of self. In D. Cicchetti & M. Beeghly (Eds.), *The self in transition: Infancy to childhood* (pp. 139-164). Chicago: Chicago University Press.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18- month-old children. *Developmental Psychology, 31*, 838–850

Meltzoff, A. N. (2007a). 'Like me': a foundation for social cognition. *Developmental Science*, *10*(1), 126-134.

Meltzoff, A. N. (2007b). The 'like me' framework for recognizing and becoming an intentional agent. *Acta Psychologica*, *124*(1), 26-43.

Miller, C. L. (1983). Developmental changes in male/female voice classification by infants. *Infant Behavior and Development, 6*, 313–330.

Misch, A., Over, H., & Carpenter, M. (2016). I won't tell: Young children show loyalty to their group by keeping group secrets. *Journal of Experimental Child Psychology*, *142*, 96-106.

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, *16*(4), 495-500.

Moya, C., & Henrich, J. (2016). Culture–gene coevolutionary psychology: Cultural learning, language, and ethnic psychology. *Current Opinion in Psychology*, *8*, 112-118.

Nielsen, M. (2006). Copying actions and copying outcomes: social learning through the second year. *Developmental Psychology*, *42*(3), 555.

Nielsen, M. (2018). The social glue of cumulative culture and ritual behavior. *Child Development Perspectives*, *12*(4), 264-268.

Nielsen, M. (2012). Imitation, pretend play, and childhood: Essential elements in the evolution of human culture?. *Journal of Comparative Psychology*, *126*(2), 170.

Nielsen, M., & Blank, C. (2011). Imitation in young children: When who gets copied is more important than what gets copied. *Developmental Psychology*, *47*(4), 1050.

Nielsen, M., Kapitány, R., & Elkins, R. (2015). The perpetuation of ritualistic actions as revealed by young children's transmission of normative behavior. *Evolution and Human Behavior*, *36*(3), 191-198.

Nielsen, M., Moore, C., & Mohamedally, J. (2012). Young children overimitate in third-party contexts. *Journal of Experimental Child Psychology*, *112*(1), 73-83.

Nielsen, M., Slaughter, V., & Dissanayake, C. (2013). Object-directed imitation in children with high-functioning autism: Testing the social motivation hypothesis. *Autism Research*, *6*(1), 23-32.

Nielsen, M., Tomaselli, K., & Kapitány, R. (2018). The influence of goal demotion on children's reproduction of ritual behavior. *Evolution and Human Behavior*, *39*(3), 343-348.

Oláh, K., Elekes, F., Bródy, G., & Király, I. (2014). Social category formation is induced by cues of sharing knowledge in young children. *PLoS One*, *9*(7), e101680.

Oláh, K., Elekes, F., & Király, I. (2019). *Creating a Shared Representational Space.* DOI: 10.31234/osf.io/g2z74

Oláh, K., Elekes, F., Pető, R., Peres, K., & Király, I. (2016). 3-year-old children selectively generalize object functions following a demonstration from a linguistic in-group member: evidence from the phenomenon of scale error. *Frontiers in Psychology*, *7*, 963.

Oostenbroek, J., & Over, H. (2015). Young children contrast their behavior to that of out-group members. *Journal of Experimental Child Psychology*, *139*, 234-241.

Over, H., & Carpenter, M. (2009). Priming third-party ostracism increases affiliative imitation in children. *Developmental Science*, *12*(3), F1-F8.

Over, H., & Carpenter, M. (2012). Putting the social into social learning: explaining both selectivity and fidelity in children's copying behavior. *Journal of Comparative Psychology*, *126*(2), 182.

Over, H., & Carpenter, M. (2013). The social side of imitation. *Child Development Perspectives*, *7*(1), 6-11.

Over, H., Carpenter, M., Spears, R., & Gattis, M. (2013). Children selectively trust individuals who have imitated them. *Social Development*, *22*(2), 215-224.

Paukner, A., Suomi, S. J., Visalberghi, E., & Ferrari, P. F. (2009). Capuchin monkeys display affiliation toward humans who imitate them. *Science*, *325*(5942), 880-883.

Paulus, M. (2012). Action mirroring and action understanding: An ideomotor and attentional account. *Psychological Research*, *76*(6), 760-767.

Paulus, M., Hunnius, S., van Wijngaarden, C., Vrins, S., van Rooij, I., & Bekkering, H. (2011). The role of frequency information and teleological reasoning in infants' and adults' action prediction. *Developmental Psychology, 47*, 976–983.

Paulus, M., Hunnius, S., Vissers, M., & Bekkering, H. (2011a). Bridging the gap between the other and me: The functional role of motor resonance and action effects in infants' imitation. *Developmental Science*, *14*(4), 901-910.

Paulus, M., Hunnius, S., Vissers, M., & Bekkering, H. (2011b). Imitation in infancy: Rational or motor resonance?. *Child Development*, *82*(4), 1047-1057.

Phillips, B.C., Seston, R., & Kelemen, D. (2012). Learning about tool categories via eavesdropping. *Child Development, 83*(6), 2057-72.

Phillips, A. T., & Wellman, H. M. (2005). Infants' understanding of object-directed action. *Cognition*, *98*(2), 137-155.

Pietraszewski, D., & Schwartz, A. (2014a). Evidence that accent is a dimension of social categorization, not a byproduct of perceptual salience, familiarity, or ease-of-processing. *Evolution and Human Behavior*, *35*(1), 43-50.

Pietraszewski, D., & Schwartz, A. (2014b). Evidence that accent is a dedicated dimension of social categorization, not a byproduct of coalitional categorization. *Evolution and Human Behavior*, *35*(1), 51-57.

Pinkham, A. M., & Jaswal, V. K. (2011). Watch and learn? Infants privilege efficiency over pedagogy during imitative learning. *Infancy*, *16*(5), 535-544.

Plötner, M., Over, H., Carpenter, M., & Tomasello, M. (2015). The effects of collaboration and minimal-group membership on children's prosocial behavior, liking, affiliation, and trust. *Journal of Experimental Child Psychology*, *139*, 161-173.

Poulin-Dubois, D., Brooker, I., & Polonia, A. (2011). Infants prefer to imitate a reliable person. *Infant Behavior and Development*, *34*(2), 303-309.

Powell, L. J., Schachner, A., & Spelke, E. S. (2014, July). *Infants' Generalization of Causal and Non-causal Actions Across Social Groups.* Poster presented at International Conference on Infant Studies, Berlin, Germany

Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. *Proceedings of the National Academy of Sciences*, *110*(41), E3965-E3972.

Powell, L. J., & Spelke, E. S. (2018a). Human infants' understanding of social imitation: Inferences of affiliation from third party observations. *Cognition*, *170*, 31-48.

Powell, L. J., & Spelke, E. S. (2018b). Third-party preferences for imitators in preverbal infants. *Open Mind*, *2*(2), 61-71.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*(2), 129-154.

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception*, *31*(9), 1109-1121.

Rakoczy, H., Hamann, K., Warneken, F., & Tomasello, M. (2010). Bigger knows better: Young children selectively learn rule games from adults rather than from peers. *British Journal of Developmental Psychology*, *28*(4), 785-798.

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: young children's awareness of the normative structure of games. *Developmental Psychology*, *44*(3), 875.

Rappaport, R. A. (1999). *Ritual and religion in the making of humanity*. Cambridge, MA: Cambridge University Press.

Renno, M. P., & Shutts, K. (2015). Children’s social category-based giving and its correlates: Expectations and preferences. *Developmental Psychology*, *51*(4), 533.

Rhodes, M. (2013). How two intuitive theories shape the development of social categorization. *Child Development Perspectives*, *7*(1), 12-16.

Rhodes, M., & Chalik, L. (2013). Social categories as markers of intrinsic interpersonal obligations. *Psychological Science*, *24*(6), 999-1006.

Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology*, *59*(3), 244-274.

Rhodes, M., Leslie, S. J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences*, *109*(34), 13526-13531.

Richter, N., Over, H., & Dunham, Y. (2016). The effects of minimal group membership on young preschoolers’ social preferences, estimates of similarity, and behavioral attribution. *Collabra, 2*, 1-8.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, *27*, 169-192.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, *2*(9), 661-670.

Rochat, M. J., Serra, E., Fadiga, L., & Gallese, V. (2008). The evolution of social cognition: Goal familiarity shapes monkeys' action understanding. *Current Biology*, *18*(3), 227-232.

Roder, B. J., Bushnell, E. W., & Sasseville, A. M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy*, *1*(4), 491-507.

Royka, A., Aboody, R., & Jara-Ettinger, J. (2018). Movement as a message: inferring communicative intent from actions. In Proceedings of the 40th Annual Conference of the Cognitive Science Society.

Rutland, A. (1999). The development of national prejudice, in-group favouritism and self-stereotypes in British children. *British Journal of Social Psychology*, *38*(1), 55-70.

Schachner, A., & Carey, S. (2013). Reasoning about 'irrational'actions: When intentional movements cannot be explained, the movements themselves are seen as the goal. *Cognition*, *129*(2), 309-327.

Schleihauf, H., Graetz, S., Pauen, S., & Hoehl, S. (2018). Contrasting social and cognitive accounts on overimitation: The role of causal transparency and prior experiences. *Child Development*, *89*(3), 1039-1055.

Schleihauf, H., Pauen, S., & Hoehl, S. (2019). Minimal group formation influences on over-imitation. *Cognitive Development*, *50*, 222-236.

Schmidt, M. F., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition*, *124*(3), 325-333.

Schmidt, M. F., Rakoczy, H., & Tomasello, M. (2019). Eighteen-Month-Old Infants Correct Non-Conforming Actions by Others. *Infancy*, *24*(4), 613-635.

Schwier, C., Van Maanen, C., Carpenter, M., & Tomasello, M. (2006). Rational imitation in 12-month-old infants. *Infancy*, *10*(3), 303-311.

Scott, R. M., & Baillargeon, R. (2013). Do infants really expect agents to act efficiently? A critical test of the rationality principle. *Psychological Science*, *24*(4), 466-474.

Shimizu, Y. A., & Johnson, S. C. (2004). Infants' attribution of a goal to a morphologically unfamiliar agent. *Developmental Science*, *7*(4), 425-430.

Shimpi, P. M., Akhtar, N., & Moore, C. (2013). Toddlers' imitative learning in interactive and observational contexts: The role of age and familiarity of the model. *Journal of Experimental Child Psychology*, *116*(2), 309-323.

Shutts, K., Banaji, M. R., & Spelke, E. S. (2010). Social categories guide young children's preferences for novel objects. *Developmental Science*, *13*(4), 599-610.

Shutts, K., Roben, C. K. P., & Spelke, E. S. (2013). Children's use of social categories in thinking about people and social relationships. *Journal of Cognition and Development*, *14*(1), 35-62.

Shutts, K., Kinzler, K. D., McKee, C. B., & Spelke, E. S. (2009). Social information guides infants' selection of foods. *Journal of Cognition and Development*, *10*(1-2), 1-17.

Skerry, A. E., Carey, S. E., & Spelke, E. S. (2013). First-person action experience reveals sensitivity to action efficiency in prereaching infants. *Proceedings of the National Academy of Sciences*, *110*(46), 18728-18733.

Sodian, B., Schoeppner, B., & Metz, U. (2004). Do infants apply the principle of rational action to human agents?. *Infant Behavior and Development*, *27*(1), 31-41.

Soley, G. (2019). What Do Group Members Share? The Privileged Status of Cultural Knowledge for Children. *Cognitive Science*, *43*(10), e12786.

Soley, G., & Aldan, P. (2020). Children and adults selectively attribute shared cultural knowledge to speakers of the same language. *Child Development*, *91*(1), e218-e230.

Soley, G., & Sebastián-Gallés, N. (2015). Infants prefer tunes previously introduced by speakers of their native language. *Child Development*, *86*(6), 1685-1692.

Sommerville, J. A., & Woodward, A. L. (2005a). Pulling out the intentional structure of action: the relation between action processing and action production in infancy. *Cognition*, *95*(1), 1-30.

Sommerville, J. A., & Woodward, A. L. (2005b). Infants' sensitivity to the causal features of means-end support sequences in action and perception. *Infancy*, *8*(2), 119-145.

Sosis, R. (2000). Costly signaling and torch fishing on Ifaluk Atoll. *Evolution and Human Behavior*, *21*(4), 223-244.

Sosis, R. (2004). The adaptive value of religious ritual: Rituals promote group cohesion by requiring members to engage in behavior that is too costly to fake. *American Scientist*, *92*(2), 166-172.

Sosis, R. (2005). Does religion promote trust?: The role of signaling, reputation, and punishment. *Interdisciplinary Journal of Research on Religion*, *1*.

Sosis, R., & Alcorta, C. (2003). Signaling, solidarity, and the sacred: The evolution of religious behavior. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, *12*(6), 264-274.

Sosis, R., & Bressler, E. R. (2003). Cooperation and commune longevity: A test of the costly signaling theory of religion. *Cross-cultural Research*, *37*(2), 211-239.

Southgate, V., Chevallier, C., & Csibra, G. (2009). Sensitivity to communicative relevance tells young children what to imitate. *Developmental Science*, *12*(6), 1013-1019.

Southgate, V., & Csibra, G. (2009). Inferring the outcome of an ongoing novel action at 13 months. *Developmental Psychology*, *45*(6), 1794.

Southgate, V., Johnson, M. H., & Csibra, G. (2008). Infants attribute goals even to biomechanically impossible actions. *Cognition*, *107*(3), 1059-1069.

Sparks, E., Schinkel, M. G., & Moore, C. (2017). Affiliation affects generosity in young children: The roles of minimal group membership and shared interests. *Journal of Experimental Child Psychology*, *159*, 242-262.

Spelke, E. S. (2016). Core knowledge and conceptual change: A perspective on social cognition. In Barner, D., & Baron, A. S. (Eds.), *Core knowledge and conceptual change* (pp. 279-300). Oxford University Press.

Spelke, E., Bernier, E. P., & Skerry, A. E. (2013). Core social cognition. *In Navigating the social world. What infants, children, and other species can teach u*s (pp. 11–16). Oxford: Oxford University Press.

Sperber, D. 1996. *Explaining Culture: A Naturalistic Approach*. Oxford: Blackwell.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.

Sperber, D., & Wilson, D. (1987). Précis of relevance: Communication and cognition. *Behavioral and brain sciences*, *10*(4), 697-710.

Staal,F. (1975). The meaninglessness of ritual. *Numen, 26*, 2-22.

Stenberg, G. (2013). Do 12-month-old infants trust a competent adult?. *Infancy*, *18*(5), 873- 904.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223*, 96–102.

Tajfel, H. (1974). Social identity and intergroup behaviour. *Information (International Social Science Council)*, *13*(2), 65-93.

Tajfel, H. (1978). Social categorization, social identity and social comparison. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the social psychology of intergroup relations* (pp. 61–76). London: Academic Press.

Tajfel, H. (1981) *Human groups and social categories: Studies in social psychology*. Cambridge, England: Cambridge University Press.

Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2405-2415.

Ting, F., He, Z., & Baillargeon, R. (2019). Toddlers and infants expect individuals to refrain from helping an ingroup victim's aggressor. *Proceedings of the National Academy of Sciences*, *116*(13), 6025-6034.

Tomasello, M. (1996). Do apes ape? In B. G. Galef Jr. and C. M. Heyes (Eds.), *Social learning in animals: The roots of culture*, (pp. 319–346). New York: Academic Press.

Tomasello, M. (1999a). *The Cultural Origins of Human Cognition*. Cambridge: Harvard University Press.

Tomasello, M. (1999b). The human adaptation for culture. *Annual Review of Anthropology*, *28*(1), 509-529.

Tomasello, M. (2019). *Becoming human: A theory of ontogeny*. Cambridge, MA: The Belknap Press of Harvard University Press.

Tomasello, M., & Barton, M. E. (1994). Learning words in nonostensive contexts. *Developmental Psychology*, *30*(5), 639.

Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, *20*(5), 454-463.

Uller, C. (2004). Disposition to recognize goals in infant chimpanzees. *Animal Cognition*, *7*(3), 154-161.

Uzgiris, I. C. (1981). Two functions of imitation during infancy. *International Journal of Behavioral Development*, *4*(1), 1-12.

van Elk, M., van Schie, H. T., Hunnius, S., Vesper, C., & Bekkering, H. (2008). You'll never crawl alone: neurophysiological evidence for experience-dependent motor resonance in infancy. *Neuroimage*, *43*(4), 808-814.

VanderBorght, M., & Jaswal, V. K. (2009). Who knows best? Preschoolers sometimes prefer child informants over adult informants. *Infant and Child Development: An International Journal of Research and Practice*, *18*(1), 61-71.

Vorobyova, L., Teglas, E., & Gergely, G. (2017, January). *Infants' understanding of multiple agents' goal-pursuit in cooperative vs competitive context.* Poster presented at Budapest CEU Conference on Cognitive Development, Budapest, Hungary.

Wagner, L., & Carey, S. (2005). 12-month-old infants represent probable endings of motion events. *Infancy*, *7*(1), 73-83.

Wang, Y., & Hamilton, A. F. D. C. (2012). Social top-down response modulation (STORM): a model of the control of mimicry in social interaction. *Frontiers in Human Neuroscience*, *6*, 153.

Watson-Jones, R. E., & Legare, C. H. (2016). The social functions of group rituals. *Current Directions in Psychological Science*, *25*(1), 42-46.

Watson-Jones, R. E., Legare, C. H., Whitehouse, H., & Clegg, J. M. (2014). Task-specific effects of ostracism on imitative fidelity in early childhood. *Evolution and Human Behavior*, *35*(3), 204-210.

Wen, N. J., Herrmann, P. A., & Legare, C. H. (2016). Ritual increases children's affiliation with in-group members. *Evolution and Human Behavior*, *37*(1), 54-60.

Wen, N. J., Willard, A. K., Caughy, M., & Legare, C. H. (in press) Watch me, watch you: Ritual participation increases in-group displays and out-group monitoring in children. *Philosophical Transactions of the Royal Society B*, *1001*, 48109-1043.

Whitehouse, H. (2012). Explaining ritual. In G. W. Dawes and J. Maclaurin (Eds.), *A new science of religion*, (pp. 81-100). New York: Routledge.

Whitehouse, H., & Lanman, J. A. (2014). The Ties That Bind Us: Ritual, Fusion, and Identification. *Current Anthropology, 55*(6), 674-695.

Whiten, A., Flynn, E., Brown, K., & Lee, T. (2006). Imitation of hierarchical action structure by young children. *Developmental Science*, *9*(6), 574-582.

Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2417-2428.

Wild, K. S., Poliakoff, E., Jerrison, A., & Gowen, E. (2012). Goal-directed and goal-less imitation in autism spectrum disorder. *Journal of Autism and Developmental disorders*, *42*(8), 1739-1749.

Wilks, M., Kapitány, R., & Nielsen, M. (2016). Preschool children's learning proclivities: When the ritual stance trumps the instrumental stance. *British Journal ofDevelopmental Psychology*, *34*(3), 402-414.

Williams, J. H., Whiten, A., & Singh, T. (2004). A systematic review of action imitation in autistic spectrum disorder. *Journal of Autism and Developmental Disorders*, *34*(3), 285-299.

Williamson, R. A., & Markman, E. M. (2006). Precision of imitation as a function of preschoolers' understanding of the goal of the demonstration. *Developmental Psychology*, *42*(4), 723.

Wilson, D., & Sperber, D. (2003). Relevance theory. In L. Horn, & G. Ward (Eds.), *Handbook of pragmatics* (pp. 607–632). Oxford: Blackwell

Wohlschläger, A., Gattis, M., & Bekkering, H. (2003). Action generation and action perception in imitation: an instance of the ideomotor principle. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 501-515.

Wood, L. A., Kendal, R. L., & Flynn, E. G. (2012). Context-dependent model-based biases in cultural transmission: children's imitation is affected by model age over model knowledge state. *Evolution and Human Behavior*, *33*(4), 387-394.

Zmyj, N., & Buttelmann, D. (2014). An integrative model of rational imitation in infancy. *Infant Behavior and Development*, *37*(1), 21-28.

Zmyj, N., Buttelmann, D., Carpenter, M., & Daum, M. M. (2010). The reliability of a model influences 14-month-olds' imitation. *Journal of Experimental Child Psychology*, *106*(4), 208-220.

Zmyj, N., Daum, M. M., & Aschersleben, G. (2009). The development of rational imitation in 9-and 12-month-old infants. *Infancy*, *14*(1), 131-141.

Zmyj, N., Daum, M. M., Prinz, W., Nielsen, M., & Aschersleben, G. (2012). Fourteen-month-olds' imitation of differently aged models. *Infant and Child Development*, *21*(3), 250-266.

ACKNOWLEDGEMENT

TO EXTERNAL FUNDING AGENCIES
CONTRIBUTING TO PHD DISSERTATIONS

Name of doctoral candidate: Nazlı Altınok

Title of dissertation: This is how we do it: Interpreting sub-efficient means actions of group members in infancy and childhood

Name of the primary supervisor: György Gergely

External funding agency: **European Research Council**

Acknowledgement: This research has received funding from the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/10254, 2015-2021) under **grant agreement No. 609819 – SOMICS**.