Author: Verschoor, Stephan Alexander
Title: Learning about goals : development of action perception and action control
Issue Date: 2014-06-25

Learning about goals:

Development of action perception and action control

By Stephan Alexander Verschoor

The research in this thesis was supported by the Netherlands Organization for Scientific Research (NWO), grant number 400-05-112.

Learning about goals:

Development of action perception and action control

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus Prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 25 juni 2014
klokke 10.00 uur

door

Stephan Alexander Verschoor

geboren te Alphen aan den Rijn
in 1975

Promotiecommissie

Promotor: Prof. Dr. B. Hommel

Co-promotor: Dr. S. Biro

Manuscriptcommissie: Prof. Dr. E. Crone
Prof. Dr. M. Paulus (LMU Munich, Germany)
Prof. Dr. N. Schiller

Content

Chapter 1: Introduction

"For the things we have to learn before we can do, we learn by doing"

(Aristotle, Nicomachean Ethics)

"We must perceive in order to move, but we must also move in order to perceive"

(Gibson, 1979)

Chapter 1

This thesis aims to answer some questions with regard to how infants learn to perceive and perform goal-directed action. To address these two interrelated aspects of goal-directed action, the thesis employs a range of behavioral, occulomotor and pupillary measures designed to tap into developmental changes and the cognitive mechanisms that underlie these abilities. The first section of chapter 1 will introduce the general topic. The second section will focus on how infants perceive third-person goal directed action. A third section will focus on how infants learn about goals by experience. The fourth section will give a short overview of the remaining chapters. Readers should expect a certain amount of overlap between chapter 1 and chapters 2 to 8 since they are submitted/ accepted/ published as stand-alone articles.

General Background

Let us consider why our question "how do infants learn to perceive, and perform goal-directed action?" is of serious importance. The world today is (as it has always been) an ever changing hostile and complex place in which humans aspire to live comfortable and fulfilling lives. Fortunately humans are endowed with the ability to act on their environment, since sitting back and letting it all flow by is not an option. In order to act effectively upon our environment we need to be able to set goals. "How are we able to set such goals?" thus can be considered a life and death question.

To start answering this question, we will reflect on what it is that constitutes a goal and a goal-directed action. A goal is a desired, future state of affairs toward which effort is directed to attain it. A goal-directed action is an action directed toward attaining a goal. These definitions imply volition (or intentionality) and a cost, since (limited) physical and/or mental effort needs to be exerted to perform the action to achieve the desired state. The definition also implies prospective memory or planning since the effort exerted needs to be directed towards changing the future environment in a meaningful way so that the goal is actually attained.

Some authors have suggested that the rudimentary capability to have first-person goals (Rizzolatti & Craighero, 2004; Rochat, 2001) and perceive third-person goals (Meltzoff & Moore, 1977) is with us from birth on—or even earlier; while others stress that the capacity is learned (e.g., Piaget, 1962). However, given the complexity and flexibility of human behavior and the dramatic changes the mind and body undergo during development, it is rather unlikely that such innate goals remain stable across the lifespan (Elsner & Hommel, 2004). Since our survival and life fulfillment critically depend on having complex goals, it seems necessary to assume that we learn new goals or modify existing goals during our lifetime.

A new question thus arises; if goals are not (all) innate, how do we acquire them? In principle the (folk-psychological) answer to that question seems quite simple, namely through experience and observing others. However, under closer, more scientific scrutiny these answers are not so straightforward. These answers lead to new questions that result in more complicated answers: “How does one acquire new goals through observation?” and “How does one acquire new goals through experience?”. These more philosophical questions inevitably result in more elaborate and complicated answers. Along these lines many authors have focused on observation of third-person action as an important source for acquiring new goals, while other theorists have focused more on acquiring new goals through active exploration. Although these questions and approaches are to some extent separated throughout the dissertation for presentational purposes, one should keep in mind that theoretically and empirically these approaches are heavily interrelated. The crosstalk between the observation of third-person action and action experience is most prominently displayed in chapter 4 wherein action experience actually alters and aids third-person action perception.

Under the assumption that new goals are acquired during the lifetime, a logical step in studying these questions is to study the emergence of goals in infants. It seems sensible to speculate that the uptake of new goals is led by the profound developmental changes in mind and body that take place during this period in life. Indeed, the concept of goal itself might develop from a primitive to a more complex concept during this period as indicated by the emergence of more complex goal-directed behaviors with progressing age. Such developmental conceptual changes in goals may be mediated by advances in short- and long-term memory (for a review, see Rovee-Collier, 1999), knowledge of the physical world (such as object permanence (Piaget, 1936), and developmental changes in the cognitive system that mediate goal representation itself. Unraveling the answer to the question mentioned above "How do we learn about goals" is an ongoing effort in cognitive science.

The current thesis will focus on developmental research on goal perception and representation. The thesis will firstly focus on how infants learn about goals by observation and then focus on the interplay of action experience and action perception. Thereafter focus will shift toward how infants acquire new goals by experience and the cognitive representations that result from such experience.

Chapter 1

Learning about goals by perceiving third-person actions

The human race is essentially social. It relies on meaningful communication for sharing knowledge on successful survival strategies gathered throughout the generations. Meaningful communication is impossible without goal perception, the ability to perceive goals in others' behavior. This ability can thus be placed at the very core of humanity. Furthermore, inferences about others' goals, enables the prediction of behavior and thus greatly enhances our ability to react correctly and timely to others. This is of great advantage in any social situation.

Indeed our everyday experiences are heavily influenced by the ability for goal perception. Goal perception is pervasive and immediately effects our perception, memory and reasoning about any third-person actions (e.g., Woodward, Sommerville, Gerson, Henderson & Buresh, 2009). To illustrate, the propensity readily makes us interpret impoverished events, like the movement of two dimensional geometrical figures, as goal directed (e.g., Heider & Simmel, 1944; Osaka, Ikeda & Osaka, 2012). Thus it seems that our cognitive systems actively try to make sense of the world by interpreting any perceived movements in terms of intentions and goals. As a result of this propensity we do not experience life as a succession of physical movements but rather parse events in terms of more meaningful terms such as goals. Indeed, the propensity to perceive goals, also help us to control our environment by enabling us to predict and explain behavior of others.

In psychology, the ability to perceive goals and predict actions of others is often captured under the term Theory of Mind (ToM). ToM as a whole consists of three cognitive abilities that are assumed to be closely related, and sometimes even as being produced by a single underlying cognitive mechanism (Ravenscroft, 1997). These abilities are: the ability to predict human behavior in a wide range of circumstances, to attribute mental states to humans, and the ability to explain the behavior of humans in terms of their possessing mental states. Furthermore, the term also refers to the theory hypothesized to underpin such processing (Ravenscroft, 1997). Although the definition mentioned above only mentions "humans", there is ample evidence that humans also use mindreading for machines and animals. However these actors are not incorporated in the definition since insisting that mental state attributions to animals are not metaphorical is compatible with such attributions being "systematically false". Therefore, the precise extension of the ToM remains under debate. Some theorists would even contend that the same processes involved in ToM also apply to first person mindreading by assuming simulational processes.

ToM, through communication, enables typical human phenomena such as language and culture. Since these phenomena are human prerogatives, some researchers postulate intention reading ability to be restricted to the human race (for a review see: Lurz, 2011). Nonetheless, some evidence indicating intention reading in animals such as chimpanzees does exist (e.g. Premack & Woodruff, 1978; for a review see Call & Tomasello, 2008). Many of the (philosophical) arguments against interpreting these studies as evidence for animal mindreading abilities, such as posing consciousness and language ability (e.g. Davidson, 1985) or "mentalese" (a hypothetical language in which concepts and propositions are represented in the mind without words) (Fodor, 1975) as prerequisites (for a review see: Lurz, 2009), may also apply to infants. Nonetheless, it seems reasonable to assume that the intention reading ability evolved to a greater extent in the human race. Thus, since full blown ToM becomes apparent later in human development, behavior that infants show, when consistent with goal attribution, may be interpreted as evidence for an emerging ability to read intentions. Nevertheless theorists are divided with regard to how ToM develops, some see infants as actively exploring the environment to accumulate and falsify hypothesis regarding the ToM (Gopnik & Meltzoff 1997), while others see it as an ability that is to a large extent innate and matures during childhood (Scholl & Leslie 1999).

However, infants' initial state of goal perception or intention reading (like that in animals) may be qualitatively different from the adult experience that utilizes the full blown ToM. It remains a question of great interest, whether infants actually ascribe mental states, such as beliefs and desires, to others. Alternatively, they may rely on contingencies between the observable cues in the environment and target behaviors, trigger stereotypical motor patterns that are innate, or rely on mental simulation that doesn't require having mental state concepts (Andrews, 2012). Indeed, theorists have proposed that intention reading need not be mentalistic in content (Gergely, Nadasdy, Csibra, & Biro, 1995). They propose that intention attribution in the mentalistic sense evolves from teleological interpretations. To illustrate: the teleological answer to the question “Why did the chicken cross the road?” would be: “To *get* to the other side” whereas the mentalistic answer would be “Because it *wanted* to go to the other side”. Notice that both answers explain the observed behavior, but only the mentalistic answer refers to the mind of the chicken. An even deeper philosophical divide concerns the subjectively perceived causal structure of mental states. On the one hand, theorists such as Fodor (1987) maintain that intentional states are in fact causal in producing behavior, while other theorists such as Dennett (1987) propose that concepts of intentional states, such as belief, desire, and perceiving, are theoretical concepts that exist by virtue of, and are determined by, a common-sense psychology or folk-psychology (ToM). According to Dennet, for another to have intentional states, is for its behaviors to be well predicted and explained by the principles of folk psychology (ToM) (Lurz, 2009). The ascribed intentional states do not necessarily refer to real structural or functional states in the brain or body. He thus proposes that ToM is epiphenomenal to behavior predictive strategies.

These philosophical considerations aside, theorists like Sodian (2011) propose, analogues to the teleological view (e.g., Gergely et al., 1995), that early infants’ goal attribution relies on a “lean level of analysis” working solely on the assumption that an agent’s (intentional) action is goal oriented. Only the later representational ToM analysis of goal attribution requires the observer to represent the agents’ viewpoint, who can have desires and (false) beliefs about a certain state of the world mentally. Furthermore, Sodian (2011) goes on to assume that the ToM analysis builds on this earlier “lean level of analysis" thus suggesting a developmental pathway. She bases this assumption on a number of studies that show that individual differences in goal attribution in habituation tasks at the ages of 6–14 months, predicted the performance on a ToM test at preschool age in several independent studies (Aschersleben, Hofer, & Jovanovic, 2008; Wellman, Lopez-Duran, LaBounty, & Hamilton, 2008; Wellman, Phillips, Dunphy-Lelii, & Lalonde, 2004). Whatever the exact ontogeny and content of goal attributions, let us now turn to methods for investigating the development of goal attribution in infancy.

Methods of assessing goal perception

Alas intention reading (just as intention proper) cannot be directly observed. This makes it imperative for developmental cognitive research on goal attribution to rely on indirect measures. Until recently most research found such a measure in imitation. The logic of these studies is that in order to successfully imitate the goal of a perceived third-person action, the participants must perceive it as goal directed. This logic is temptingly simple. Nonetheless, it becomes rather complicated when one takes into account the concept of emulation (Tomasello, Davis-Dasilva, Camak, & Bard, 1987). This idea holds that one can copy behavior, outcomes of behavior or learn affordances without actually attributing a goal to the observed model. Thus the exact definition of imitation adhered to is of great importance for research on goal perception. Some have suggested definitions that require imitation to refer to the learning of novel actions (Byrne & Russon, 1998). Others suggested that the term should involve intention understanding (Tomassello, 1999) or a conscious intention to imitate (Tissaw, 2007). Still others see these strict definitions as too controversial (Paulus, 2011) and adhere to a definition that only requires the actions of the imitator to be sufficiently similar to the model in terms of behavior and outcome (for a review see Paulus, 2011). However, imitational studies, adhering to different levels of strictness on the definition, have been used extensively to tap into a wide variety of cognitive abilities, including learning new goals by observation, even though relatively few experiments are devoted to understanding the phenomenon of imitation itself (Jones, 2009).

One of the aims of Piaget's elaborate genetic epistemological stage theory (1962) was to account for the emergence of imitation. Others have suggested an innate apprehension of equivalences between acts of the self and those of others that enables imitation (Melzoff, 2007). Meltzoff & Moore (e.g., 1977) and others have presented evidence suggesting that a rudimentary innate ability to imitate (very simple) gestures such as protruding ones tongue or mouth opening is present in neonates (see Anisfeld, 1991 for a review). These findings are under heavy debate (Jones, 2009) and even if one does accept the behavior in these studies as imitation or emulation of a goal-directed action, the goal itself is not new in the sense that the infants never protruded their tongues before. Thus although imitation is seen as one of the most important steps in the development of the human brain (Hayek, 1952), its onset and the exact cognitive mechanisms that enable the translation of third-person goal-directed behavior into first person behavior remain open for discussion (Jones, 2009).

Imitational research has provided a wealth of concepts regarding action perception (e.g., Meltzoff, 2007), nonetheless one could conclude that methods that use imitation as a derived measure of goal attribution might not be ideally suited to the task. There are several reasons for this conclusion. Firstly, the concept of imitation itself is ill defined. Furthermore (partly due to its troublesome definition) its developmental onset is vague at best. Differences in the onset of goal perception versus imitation could thus produce results that under- or overestimate the capacity of goal perception. Additionally, there is a large motivational component in imitational studies, meaning that infants might simply not “want” to imitate a certain action while being perfectly able to do so. Lastly, infants might “want” to imitate a certain action without having proper motor control to do so. Thus infants may be more, or less able to perceive goals than their imitation behavior suggests.

A method that may be more suitable for measuring goal perception in third-person action is the preferential looking time method as developed by Fantz (1958). In (simple) preferential looking time studies infants are habituated to (or familiarized with) a certain stimulus until attention decreases by a predefined amount. Thereafter they are confronted with a second stimulus that is different on some dimension. Attention recovery or looking longer at the second stimulus is taken as evidence that the infants noticed the change along that dimension and thus perceived the dimension, which resulted in more elaborate processing. If applied meticulously and with the right control conditions, this technique could in principle asses if and how infants perceive an action as goal-directed if one changes the stimuli in the right way. Although, just like in imitation, the cognitive mechanisms subservient to the phenomenon of habituation are under discussion (Colombo & Mitchell, 2009), preferential looking time has an important advantages over imitational measures; since the phenomenon is robustly present at birth (e.g., Friedman, Nagy & Carpenter, 1970) the measure does not run the risk of failing to detecting goal perception due to onset issues. The vast majority of recent studies on goal perception utilize this method. Recent advances in the suitability for infant research of eye tracking, EEG and near infrared spectroscopy have brought these methods within the grasp of developmental cognitive research.

Some seminal findings using looking time measures

A seminal way of assessing goal attribution in infants, that uses looking time as a measure was developed by Woodward (1998). Woodward devised a (dis-) habituation paradigm to contrast the surface features of a goal-directed action to its more conceptual aspect that is the goal of the action, the Violation of Expectation (VoE) paradigm. Using a live puppet-show she habituated 5- and 9-month-old infants to an arm coming in from the side of a stage repeatedly grasping one of two different toys presented at horizontally distinct locations (a familiar action since infants start to reliably reach for interesting thing around the fifth month of age (Bertenthal & Clifton, 1998; Clearfield & Thelen, 2001), see Figure1.

Before the test trials infants were shown that the location of the two toys were swapped. Woodward then measured the dishabituation to two different test events; in one of the events the same toy was grasped at the new location and in the other test events the new toy was grasped at the old location. In this study 9-month-olds dishabituated more to the new toy same location presentation than to the old toy different location presentation. The infants now expected the actor to continue to grasp the same toy even though it was in a different location. With appropriate controls, this study was taken as evidence for goal attribution for familiar human actions in 9-month-old infants. Goal attribution is now widely accepted to emerge in infants around 6 months (e.g., Biro & Leslie, 2007; Woodward, 1998; Kamewari, Kato, Kanda, Ishiguro, & Hiraki, 2005).

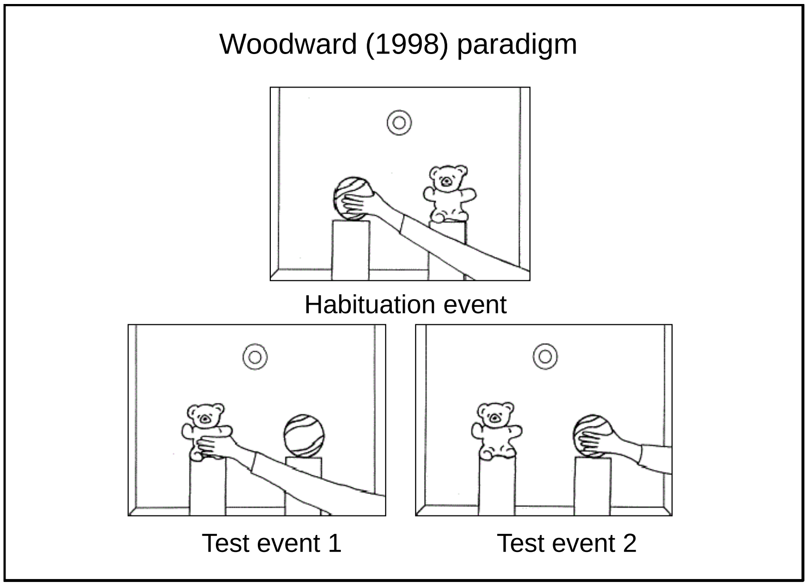

Figure 1. Woodwards 1998 paradigm. Infants were habituated to a hand grasping a toy at one of two locations. From habituation to test the toys change sides. Infants dishabituate more to a change of goal (Test event 1) than to a change of location (Test event 2) indicating the infants expects the actor to have the same goal as during habituation. This figure is partly adapted from Woodward (1998).

Another seminal line of research into goal attribution using similar methodology, aimed at assessing how infants rely on the selection of action means for goal attribution, was started by Gergely et al. (1995). In their experiment twelve-month-old infants were habituated to an animated circle that approached another circle either in an efficient or inefficient manner. In the “efficient approach” condition, the circle jumped over an obstacle in order to get to the other circle, while in the “inefficient approach” condition the circle jumped although the obstacle was not located between the two circles. In the test phase the obstacle was removed. Both groups saw the old (now inefficient) jumping action and a new straight pathway approach in which the animated circle moved towards the other circle in an efficient straight path. Gergely and colleagues (1995) found that only the infants habituated to the efficient goal approach dishabituated significantly to the inefficient, more familiar jumping action even though the efficient straight line approach was new to them. They concluded that only the infants who had seen an efficient action during habituation had seen the action as goal directed and thus had the expectation in the test phase that the circle would approach the other circle in an efficient goal-directed manner (a straight line). This result was later replicated in infants as young as six- and a half months old (Csibra, 2008).

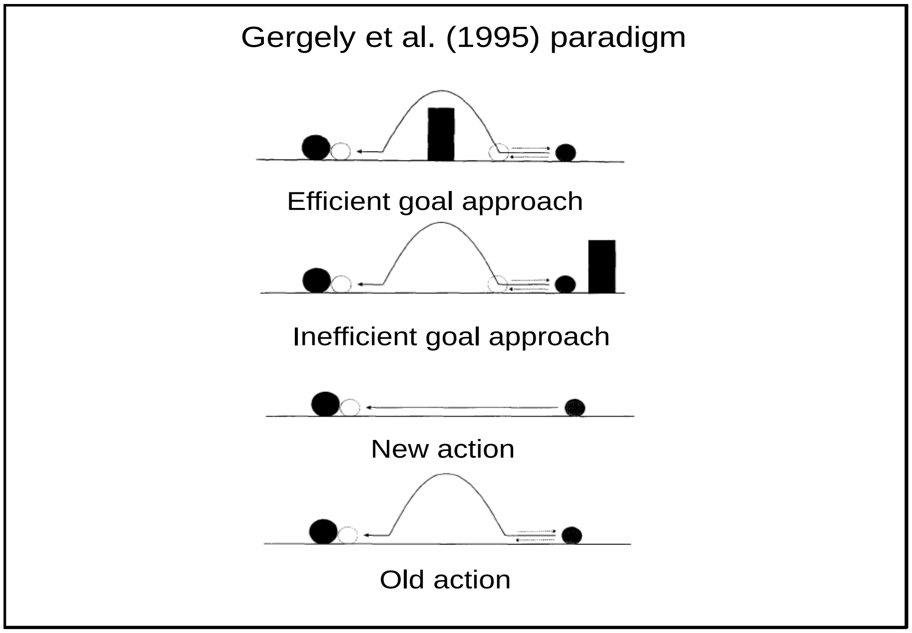

Figure 2. Gergely et al. (1995) paradigm. Infants were habituated to either the efficient or the inefficient goal approach wherein a small circle approaches a big circle (in the inefficient condition the obstacle was not in between the two circles yet the circle still jumped). During test they were shown either the old or the new action now without any obstacles. Only infants in the efficient goal approach condition dishabituate more to a the old action than to the new action. This indicates the infants expected the small ball to efficiently approach the big ball given the changed situational constraints. This figure is partly adapted from Gergely et al. (1995).

Theoretical approaches to third-person action perception

By now there is a large body of evidence that demonstrates an emerging understanding of third-person action in infancy during the second part of the first year (e.g., Biro & Leslie, 2007; Woodward, 1998; Kamewari, Kato, Kanda, Ishiguro, & Hiraki, 2005). Yet, the cognitive mechanisms that subserve such an understanding and their developmental origins are still under heavy theoretical debate. Proposed theories differ along a number of dimensions, mostly motivated by the scientific traditions from which they evolved, and thus deliver partially contrasting explanations (Paulus, 2011). However, the mechanisms suggested are often not mutually exclusive but rather should be seen as complementary, reflecting on different processes involved in goal attribution (that nonetheless might follow a certain temporal order in development). Theoretical confusion and debate is for a large part due to inaccurate definitions within the field, and the often difficult to prove claims of nativism versus constructivism or maturation versus experience.

The nativist- versus constructivist- divide in third-person goal perception, closely mimics the historic divide in language-acquisition theory; while Chomsky (1965) suggests innate modules specifically evolved for language acquisition, Vygotsky (1985) stresses experience as the main drive behind language acquisition. Similarly some propose hardwired full-fledged action understanding (Fodor, 1987) or “core principles” to guide emerging intention reading (e.g., Carey & Spelke, 1994), while others stress constructivist sensori-motor processes whereby action understanding emerges through experience (e.g., Piaget, 1962). Another dimension that differentiates theories is whether individual (e.g. Gergely & Csibra, 2003) or social processes are highlighted (Király, Csibra & Gergely, 2013).

To illustrate, several, sometimes interrelated processes have been suggested as fundamental for a full-fledged percept of a third-person goal-directed action. For one, the observant has to be able to parse the continuous stream of events in a way that enables him to select a certain sequences of events that is likely to contain a goal-directed action. Several different processes might play a role in such a parsing ability. A process hypothesized to be important for the selection of events that could contain goal-directed action, is that of actor detection. The logic behind this hypothesis is that the mere presence of a moving actor is in itself is a good predictor of goal-directed action occurring. Some theorize that the process of actor detection develops as a result of infants’ experiences with human agents (e.g., Meltzoff, 1995; Woodward, Sommerville & Guajardo, 2001) suggesting that the detection of actors first applies to humans and then generalizes to other classes of actors. Others theorize that the detection of actors relies on more abstract cues such as biological motion, non-rigid transformation and self-propulsion (e.g., Gergely et al., 1995) suggesting that the detection of actors should work for other actors besides human.

Another process, hypothesized to be important for parsing events, is the detection of significant changes in the physical world, in other words, the detection of action effects. Such effects are hypothesized to aid the interpretation of an end state as a goal state. Again, several mechanisms have been suggested for this interpretational process. Some stress the detection of surface features of the end state such as its' saliency (for a review, see: Elsner 2007) or spatiotemporal properties (e.g., Gergely et al., 1995), while others stress more elaborate cognitive processes such as the internal emulation of the action to assess the end state as a goal (Meltzoff, 2007). Another major point of divergence for theories on the perception of goal-directed action is the perception of novel, versus familiar actions. Some experience-oriented theorists, working in the grounded or embodied cognition tradition, such as Woodward (1998) stress that in order to interpret an action as goal-directed the action has to be within the infants action repertoire. Others stress mechanisms that do not rely on action repertoire but utilize more abstract interpretational principles (e.g., Gergely et al., 1995).

As outlined above theoretical standpoints clearly diverge on a number of dimensions. For clarity three of the main theories regarding third-person action understanding will now be highlighted: motor resonance, teleological reasoning and action-effect theory.

Motor resonance

A theoretical approach that has lately enjoyed major interest in the field of goal-directed action perception is the motor resonance approach. Although a privileged link between action and perception has been suggested as early as the late nineteenth century (James, 1890), approaches that emphasize such a link, the grounded- and embodied-cognition approaches (e.g., Barsalou, 2008), have recently gained renewed prominence due to the discovery of the Mirror Neuron System (MNS) (Di Pellegrino, Fadiga, Fogassi, Gallese & Rizzolatti, 1992). Both the Mirror Neuron Hypothesis- (MNH) and embodied- and grounded- cognition theories state that cognition (and thus action understanding) is grounded in bodily experiences.

Di Pellegrino and colleagues (1992) were the first to document neurons in the premotor cortex of macaque monkeys that fire when executing a grasping movement, and when a similar action is observed. This result was only obtained for goal-directed (object-directed) actions; mimicry of such actions did not produce the effect. Such findings gave rise to the idea that monkeys use the same brain regions responsible for planning actions to interpret third-person actions. Theorists suggested that by feeding information about the observed action into their own motor system, the monkeys were able to interpret the observed action in terms of a goal by simulating the action as if they were performing it themselves. Later findings suggesting a homologous MNS in humans (e.g., Gazzola & Keysers, 2009), extended this theory of action understanding through inverse planning to humans. A number of researchers went on to proposed the MNS and its' ability to "share" motor representations, to be the foundational cornerstone for higher order social processes, such as motor learning, imitation, perspective taking, understanding facial emotions and empathy (e.g., Rizzolatti, Fogassi & Gallese, 2001). Although the theory has many proponents, there are also some theorists that fear that over-interpretation of the available data gave rise to the MNH (e.g., Hickok, 2009). They propose that mirror neurons play only secondary roles in action understanding.

The most comprehensive developmental theory regarding third person action perception based on the MNH was developed by Meltzoff (2007). In his "Like me framework" he poses that infants use the (supramodal) codes generated by the MNS for interpreting third person action. Infants ascribe internal states such as goals by applying the simple strategy of attributing the same mental states that went along with the action when they perform it themselves. Although he acknowledges the capacity to emulate third-person action in the MNS develops during the first year of life he attributes this mainly to action experience. He states that a primitive, developing capacity to represent “supramodal” codes generated by the MNS is present at birth. He derives this assumption from his imitation research on neonates (Meltzoff & Moore, e.g. 1977). Others see the MNS as an emergent property of hebbian- (Hanuschkin, Ganguli & Hahnloser, 2013; Del Giudice, Manera, & Keysers, 2009) or associative- (Heyes, 2010) learning mechanisms and thus experience. Again these views are hard to disentangle experimentally. However, the main prediction of the MNH and the “Like me framework” is that the ability to perceive goal-directed action develops as a function of action experience and that it is restricted to human (or human like) agents.

The application of motor resonance theory gained large scale support in developmental science due to the seminal VoE work of Woodward (1998) (for an explanation of the methodology see above and Figure 1). Woodward found evidence for goal attribution in 9-month-old infants, but only when the puppet show was performed by a human arm, not when the actor was a rod, or when the actor performed an unfamiliar back-of-the-hand-dropping action (Woodward, 1999). This was taken as evidence that infants have to be familiar with the action presented in order to interpret it as goal directed. Furthermore it suggests that the actor has to be human (some might comment that the macaques’ MNS reacted to human motion, however humans are more like macaques than like rods). Furthermore Guajardo and Woodward (2004) showed that 7- and 12-month-old infants in a similar paradigm were unable to attribute a goal to an actor wearing a glove unless they were familiarized with the actor wearing gloves prior to the task. Sommerville, Woodward and Needham (2005) additionally showed that 3-month-old infants can attribute a goal when they observe an actor wearing a mitten if they are allowed to practice making contact and picking up toys by using a mitten covered with Velcro fabric.

Nonetheless, there is evidence suggestive of other processes playing a role. For instance several studies (e.g., Balargeon & Luo, 2005; Biro & Leslie, 2007; Luo, 2011) have found evidence of goal attribution in infants to non-human actors. These studies used paradigms very similar to that of Woodward (1998). Biro, Csibra and Gergely (2007) argue that not only cues such as human features or human-biomechanical motion can elicit identifying an actor as such, but additionally propose cues such as self propulsion and behavioral cues such as equifinal movement to account for goal attribution to non-human agents. Furthermore Luo and Baillargeon (2005) found that goal attribution in a Woodward type paradigm did not take place when during familiarization only one object was presented, suggesting that infants only attribute a goal when the actor can express a preference. Thus, although the MNH has generated research that indicates that infants more readily interpret human- as opposed to non-human action as goal directed, this may be due to additional agency cues in human action not to the MNS. Thus the MNH may yet turn out to be an ineffective “weapon of mass explanation”.

Action-Effect Learning

Another theory related to motor resonance regarding goal attribution, that also emphasizes action experience as an important factor for action perception, is ideomotor theory. The theory states that infants perceive action goals by using previously acquired action-effect representations. The nature of the representations used however is more specified than in general motor resonance theory. Repeated experience with an action and its effect automatically binds the distal representations of the action and effect into an action-effect association. These action-effect associations are bidirectional. This is crucial to the theory since this allows the agent to activate the appropriate action by thinking of its effect. Furthermore, it allows predicting action outcomes by way of activation of the motor pattern. Notice that this process can only take place if a salient effect is presented. The theory thus predicts that bidirectional action-effect associations, and therefore goal perception can only occur if a salient action effect is presented.

Although the theory was initially non developmental and mainly used to explain first-person action perception (see section 2 of the introduction), by now it has been extended to the development of third person action perception. This was done under the influence of theoretical elaborations on action-effect theory like the Theory of Event Coding (TEC) (Hommel, Müsseler, Aschersleben, & Prinz, 2001) and the discovery of the MNS. In TEC the authors proposed that humans use the same distal features to represent their own and others' actions. Although these distal codes are similar to the "supramodal" codes Meltzoff (2007) proposes, TEC does not make the assumption that they are inborn. Thus action-effect theory is in itself not sufficient to explain goal attribution in third-person action. It additionally needs some kind of internal representational overlap between first-person- and third-person action representations.

Indeed, salient action effects have been shown to aid third person goal attribution (e.g., Biro & Leslie, 2007; Jovanovic, Kiraly, Elsner, Gergely, Prinz & Aschersleben, 2007, for a review, see: Hauf, 2007) even for unfamiliar actions (Kiraly, Jovanovic, Prinz, Aschersleben, & Gergely, 2003). Furthermore infants are more likely to imitate actions that have a salient action effect (Hauf & Aschersleben, 2008; Klein, Hauf & Aschersleben, 2006; for a review, see: Elsner 2007; Melzoff, 2007). And several studies have shown that own action experience (and thus action-effect knowledge) does affect the perception of third-person goal directed action (e.g., Hauf, Aschersleben & Prinz, 2007; Sommerville & Woodward, 2005a, 2005b).

Teleological reasoning

A third important theory regarding goal attribution to observed action, that directly confronts the motor resonance based view, was first formulated by Gergely et al. (1995) (for an explanation of the methodology of their experiment see above and Figure 2). The theory belongs to the tradition that poses domain specific (possibly innate) core principles that govern developing cognition. The theory holds that infants perceive an action as goal directed when it follows the principle of rational action which states that *"an action can be explained by a goal state if, and only if, it is seen as the most justifiable action towards the goal state that is available within the constraints of reality"* (Csibra & Gergely, 1998). The action thus needs to be efficient towards the end state, considering the situational constraints, for goal attribution to occur. The theory is teleological in the sense that it does not require any mentalistic explanations to be attributed by the observer it requires that the end state of a certain action be explained by the action and the situational constraints (contrary for instance to the "like me" framework of Meltzoff (2007). Teleological reasoning is considered the core of the full blown theory of mind which allows humans to attribute intentions, feelings and (false) beliefs to others (Premack & Woodruff; 1978). Csibra and Gergely (1998) state that the principle of rational action is the initial state of the infants' naïve psychological theory whereupon can be elaborated by adding mentalistic explanations later in development**.**

The results of Gergely et al. (1995) and others who used similar paradigms (e.g., Csibra, Gergely, Biro, Koós, & Brockbank, 1999; Csibra, Biro, Koós, Gergely, 2003; Csibra, 2008; Wagner & Carey, 2005; Sodian, Schoeppner, & Metz, 2004) indeed suggest that means-selection information is crucial to goal attribution. Support for the theological stance also comes from imitation studies (e.g., Gergely, Bekkering & Kiraly, 2002) which show that infants' imitation also depends on the efficiency of the shown action (for a contradictory view on these results see Paulus, Hunnius, Vissers & Beckering, 2011a-2011b; Beisert et al., 2012).

Taken together the theoretical views outlined above provide several different mechanisms which could explain the emergence of goal perception in third-person action. In the first three empirical chapters of this thesis we will assess how each of these mechanisms affects goal perception and, where possible, try to contrast the different theoretical viewpoints.

Learning new goals by experience

The second question explored in this thesis is "How do infants acquire new goals through experience?". Infants start to show evidence of (primitive) being goal-directed in their behavior from a very early age on. The onset of intentional behavior depends on the exact definition of goal-directed action adhered to. If one does not dwell too long on in how far these behaviors are actually planned and intentional, primitive and easily disrupted goal-directed behaviors such as moving a hand towards the mouth (Butterworth & Hopkins, 1988) or orienting towards sound (Zelazo, Brody, & Chaika, 1984) are evident in neonates (e.g., Bertenthal, 1996; Metzoff & Moore, 1997; von Hofsten, 2004). Slightly more demanding behaviors such as grasping an interesting object in order to examine it, appear around 5 months of age (Bertenthal & Clifton, 1998; Clearfield & Thelen, 2001). More sophisticated intentionally planned goal-directed behavior, for which infants need to be able to distinguish means from ends, starts to appear around nine months of age (Claxton, Keen, & McCarty, 2003; Hauf, 2007; Willatts, 1999; Piaget, 1936; Woodward & Sommerville, 2000; Woodward et al., 2009).

Altogether these findings underline a developing ability to represent first-person action goals. In principle goals could be innate and just mature during development. However, given the world's ever-changing nature and complexity, and the dramatic development the mind and body of infants undergo, it is plain common sense to assume that goal representations are not innate, finite and permanent (e.g., Greenwald, 1970; Hommel & Elsner, 2009).

Piaget (1936) was the first cognitive developmental psychologist, to systematically research the origin of emerging knowledge during infancy. In his genetic epistemological approach (origin of knowledge) he firstly suggested that goals should be adaptive to the infant's changing skills and abilities. He suggested that goals may derive from its own sensorimotor exploration and experience motivated by a predisposition to adjust to its environment. Although he states that such exploratory actions are integrated with their effects into schemata necessary for perception and action, he does not elaborate on the underlying cognitive mechanism (Piaget, 1954).

Ideomotor theory

Exactly such a mechanism was proposed in the late nineteenth century, though largely neglected in psychology for a century. Following the lead of Lotze (1852) and Harless (1861), James (1890) suggested a cognitive mechanism that does what Piaget (1936) proposed; it provides actors with action goals that are rooted in their own sensorimotor experience. In his ideomotor theory James stated that all actions are necessarily involuntary when being carried out for the first time. Indeed, if one defines action as goal-directed movement, it presupposes some sort of anticipation of its effect. This again implies knowledge on action-effect relationships, which needs to be acquired before the action can be carried out “in order to” produce the outcome intentionally. Ideomotor theory suggests that such knowledge is acquired on the fly: whenever people move, they automatically and unintentionally create bidirectional associations between the perceived effects and the motor pattern producing them. This association brings the movement under voluntary control: Once acquired, the agent can now activate the motor pattern producing a movement by “thinking of” (i.e., endogenously activating the representation of) a perceptual effect. Indeed, infants start to motor babble (i.e., produce random movements) in utero (cf., Meltzoff & Moore, 1997)—which could explain the possible presence of goal representations at birth—and they are consistently exploring their environment. This provides ample opportunity to acquire movement/action-effect associations and thus a steadily increasing pool of possible action goals. Thus, James considered bidirectional movement/action-effect associations the fundamental building blocks of intentional action and provides a mechanism that could allow the emergence of goal-directed action in infants. Ideomotor theory was revived and refined by Greenwald (1970), Prinz (1990, 1997), and Hommel (1996; Elsner &

Hommel, 2001) and is now part of a broader theoretical movement stressing the interplay between perception and action (Hommel et al., 2001; Meltzoff, 2007; Meltzoff & Prinz, 2002).

Seminal findings

The quintessential paradigm that showed the formation and use of bidirectional action-effect associations was developed by Elsner and Hommel (2001). They designed the “two stage model of voluntary action” (Figure 3). Their paradigm closely resembles the two stages. In the first stage (the acquisition phase) actions are automatically bidirectionally bound to their effects. To accomplish this they had adults carry out self-chosen left and right key presses in response to a visual trigger stimulus. Each key press produced a particular sound (e.g., left key → low tone, right key → high tone), and even though these sounds were irrelevant for the task, it was assumed that participants would acquire bidirectional associations between the key presses and the tone representations. The second stage is the stage wherein a voluntary action is produced by anticipation of its action effect. Participants were again freely choosing left and right key presses, but in this test phase the visual trigger was replaced by an auditory trigger stimulus: high and low tones (identical to the previous action effects) presented in random sequence. As predicted, people were quicker and more likely to choose the action that previously had produced the currently presented trigger tone (e.g., they were quicker and more likely to press the left key when hearing the low than the high tone), suggesting that key presses and tones were indeed associated in a bidirectional fashion.

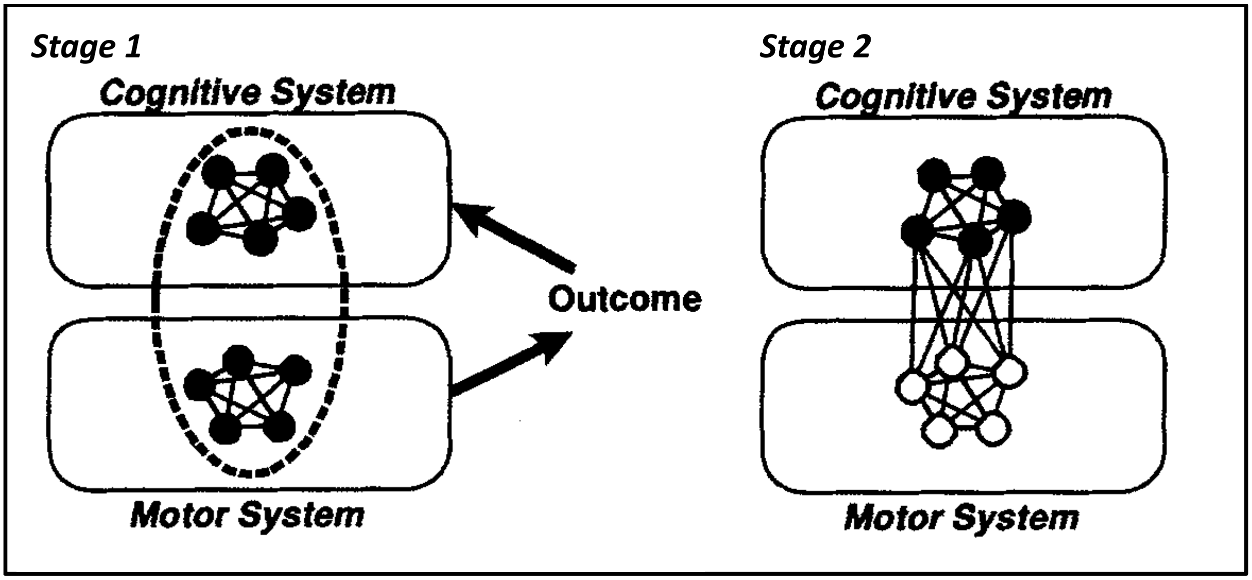

Figure 3. Two stage model by Elner and Hommel (2001). At stage1, the motor patter producing a particular effect is automatically integrated with the cognitive codes representing this effect. At stage 2, the motor pattern is intentionally executed by activating the cognitive codes that represent its effect. This figure is partly adapted from Elsner and Hommel (2001).

This seminal finding motivated numerous demonstrations of bidirectional action-effect acquisition in humans ranging from 4-year-olds (Eenshuistra, Weidema & Hommel, 2004; Kray, Eenshuistra, Kerstner, Weidema & Hommel, 2006) to adults (e.g., Elsner & Hommel, 2001). Action-effect acquisition was found for a wide range of actions and effects (for a review see: Hommel & Elsner, 2009), suggesting a general action-effect integration mechanism. Additionally, action-effect acquisition has been found after just one trial (Dutzi & Hommel, 2009), suggesting that the mechanism is fast-acting and implicit. Action-effect acquisition is modulated by the same factors that influence instrumental learning (e.g., temporal contiguity and contingency of movement and effect: Elsner & Hommel, 2004) and does not depend on voluntary attention (Dutzi & Hommel, 2009; Elsner & Hommel, 2001; Band, Steenbergen, Ridderinkhof, Falkenstein & Hommel, 2009). Together with the fact that it was also found in animals (see Elsner & Hommel, 2001), this suggests that action-effect integration it is a fairly low-level and automatic process (Elsner & Hommel, 2004). Except for Band et al. (2009) the studied mentioned above all involve the activation of actions thru presenting the effect that previously produced them. However recently a new type of paradigm emerged that assesses activation of expected effects by producing the action that normally precedes them. In such paradigms sensory attenuation (attenuation, in terms both of phenomenology and cortical response) to expected action effects is found (e.g., Cardoso-Leite, Mamassian, Schütz-Bosbach & Waszak, 2010; Hughes, Desantis, & Waszak, 2013b; for a review see: Hughes, Desantis & Waszak, 2013a).

Corroborating evidence for ideomotor processes in infancy

If bidirectional action-effect associations are indeed the fundamental building blocks for intentional action, the system that generates these associations should be operative early in life. Especially since infants show evidence of goal-directed behavior from a very early age on. As discussed above, action-effect knowledge has been implicated to be operational in higher order cognitive functions such as action understanding in 7 months-olds (e.g., Biro & Leslie, 2007; for a review, see: Hauf, 2007; Kiraly et al., 2003) and imitation in 9-months-olds (Hauf & Aschersleben, 2008; Klein et al., 2006; for a review, see: Elsner 2007; Meltzoff, 2007). Even though these findings do not provide direct evidence for bidirectional action-effect acquisition, theories that emphasize similar representational formats for first-person experience and observed action (e.g., Fabbri-Destro & Rizzolatti, 2008; Hommel et al., 2001; Meltzoff, 2007; Tomasello, 1999), and conceptualize action understanding as inverse planning (Melzoff, 2007; Baker, Saxe & Tenenbaum, 2009) consider them corroborative. Other corroborating evidence was found in studies that show very young infants to be sensitive to action-effect contingencies. For instance, newborns actively adjust their sucking rate in response to their mothers' voice as ongoing conditional feedback (DeCasper & Fifer, 1980) and 2-month-olds pursue interesting action effects by intentionally varying their sucking rate (Rochat & Striano, 1999) or varying gaze direction (Watson, 1967; for a review, see Gergely & Watson, 1999). Another line of research by Carolyn Rovee-Collier shows that action effects aid memory retrieval for actions from two months of age (Rovee & Rovee, 1969; for a review, see Rovee-Collier, 1999). Telling as these studies may be (they show that action contingent effects play an important role in infant behaviour and memory) they were not designed to directly asses the bidirectionality of action-effect associations and their use for action planning and may thus confound actual action-effect learning with simple operant conditioning. Nonetheless, these findings may reflect a developing ability for learning action-effect contingencies.

Validating ideomotor theory as a developmental theory

Ideomotor theory is rapidly gaining a following among developmental scientists. Many researchers apriori assume that learning new goals (learning new bidirectional action-effect associations) should be supported by the same mechanism as it is in adults. Although there is a lot of corroborating evidence suggestive of ideomotor processes in infants, the conclusion that the same mechanisms operates in infancy is premature since the most important piece of the puzzle is still missing: direct evidence for uptake of bidirectional action-effect associations in infancy. Such evidence would greatly validate the application of ideomotor theory as a developmental theory for goal perception. Furthermore, it would greatly broaden the scope for ideomotor theory as a whole. Thus the first aim of the second section of the thesis will be to generate such evidence. Under the assumption that we will succeed at fulfilling the first aim, the second aim regarding the application of ideomotor to developmental psychology will be to discover the developmental pathway the uptake of bidirectional action-effect associations follows.

The first crucial challenge in our quest to find direct evidence for bidirectional action-effect associations in infants will be to come up with a paradigm that is both suitable for infants and can provide such evidence. To meet this challenge we will modify the original Elsner & Hommel (2001) paradigm and rely on a number of different measures. Modifications will include great simplification of the paradigm and making the paradigm attractive to infants. Furthermore, we will employ the relatively new technique of eye tracking to overcome any motoric challenges infants face. Lastly we will use the recently rediscovered technique of pupillometry.

Chapter 1

Outline of the thesis

As mentioned above the main question of the current thesis can be divided into two, to a certain degree, interrelated questions: “How do infants learn about goals be perceiving third-person action?” and “How do infants learn new goals by experience?”. The questions of the first part of the thesis, regarding the perception of third-person action, will mainly concern what kinds of information infants use for goal attribution, and their relative importance. The importance of the chosen outcome of an action will be compared to the importance of the efficiency of the action in regard to the (changing) situational constraints. Furthermore, we will explore the role action experience plays in goal attribution. Another aim of this part of the thesis will be to establish if the different VoE paradigms (one type of paradigm concentrates on outcome selection information, the other type systematically varies the efficiency of actions in relation to changing situational constraints) tap into a unitary concept of goal. The questions of the second part of the thesis, regarding the way infants pick up new goals by experience, will mainly concern the conformation of the ideomotor processes for first-person action perception and action control in infancy. In addition, the developmental timeline of these principles will be investigated.

Chapter 2 investigates the relationship between findings from the Woodward-(1998) and the Gergely et al. (1995) type paradigm. In the Woodward-type paradigm infants are shown actions in which the actor consistently chooses one of two possible outcomes, whereas in the Gergely-type paradigm the actors use efficient means, depending on situational constraints, toward one outcome. Since these paradigms capitalize on different kinds of information the question arises which type of information is more important for goal attribution. To answer this question, we combined the two types of information into one VoE paradigm.

Chapter 3 utilizes a similar paradigm to investigate if infants can transfer goal attribution from a situation in which only means information is presented toward one goal, to a situation in which there two possible outcomes. In other words, does observing an actor efficiently adjusting his action to situational constraints, lead infants to expect that the actor will continue to approach the same goal after another potential outcome is presented in the scene? Such transfer would indicate that both the Woodward- (1998) and Gergely et al.-(1995) type paradigms tap into the same unitary concept of goal that relies on both outcome- and means- information.

In the fourth chapter we explore the question if first-person experience with a certain movement, or only first-person action-effect associations, can influence goal attribution to a novel action wherein no effects are presented.

The second part of the thesis will concern the second of our two questions “How do infants learn new goals by experience?” To further validate ideomotor theory as a cognitive developmental theory, as a starting point, chapter five replicates previous findings (Eenshuistra et al., 2004; Kray et al., 2006) that indicate that 4-year-old children can form bidirectional action-effect associations and compare their performance to adults. The secondary aim of the study was to investigated whether action-effect associations are also acquired under explicit-learning conditions and whether familiar action-effect relations (such as between a trumpet and a trumpet sound) are learned the same way as novel, arbitrary relations are.

The next two chapters represent the first attempts to provide direct evidence for the spontaneous acquisition of bidirectional action-effect associations in infancy. In chapter six we applied a highly simplified version of the Elsner and Hommel paradigm (2001).

Chapter seven is methodologically more advanced adaptation of the Elsner and Hommel paradigm (2001). The paradigm was made suitable for all age groups ranging from 7-month-old infants to adults and employed a novel pupillometric and oculomotor paradigm to study developmental changes in the role of action-effects in the acquisition of voluntary action across the lifespan. Our findings suggest that both 7- and 12-months olds (and adults) can use acquired action-effect bindings to predict action outcomes but only 12-months-olds (and adults) showed evidence for employing action-effects to select actions. This dissociation supports the idea that infants acquire action-effect knowledge before they have developed the cognitive machinery necessary to make use of that knowledge to perform intentional actions.

In the last empirical chapter of the thesis we used the newly developed methodology from chapter seven to investigate, from an ideomotor point of view, how infants represent sequential action. We aimed to contrast chaining-, concurrent- and integrated models of sequential-action representation. 9- and 12-month olds were taught action sequences consisting of two elementary actions. Thereafter the secondary action was selectively activated to assess any interactions with the primary action.

Finally chapter 9 will summarize the most important theoretical insights the thesis has to offer and suggests future research directions.

The seven empirical chapters in this thesis are either published, under revision, or submitted in international psychological journals. They are inserted in the thesis in their original submitted or published form. To honor the co-authors, a list of references is presented:

Chapter 2,

Verschoor, S. A. & Biro, S. (2012). Primacy of information about means selection over outcome selection in goal attribution by infants. *Cognitive Science, 4*, 714-725.

Chapter 3,

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

Chapter 4

Biro, S., Verschoor, S. A., Coalter, E., & Leslie, A. M. (Under revision). Outcome potential influences twelve-month-olds' interpretation of a novel action as goal-directed.

Chapter 5

Verschoor, S.A., Eenshuistra, R., Kray, J., Biro, S. & Hommel, B. (2011). Explicit learning of arbitrary and non-arbitrary action-effect relations in adults and 4-year-olds. *Frontiers in Psychology, 2*,354.

Chapter 6

Verschoor, S. A., Weidema, M., Biro, S. & Hommel, B. (2010). Where do action goals come from? Evidence for spontaneous action-effect binding in infants. *Frontiers in Cognition, 1*,201.

Chapter 7

Verschoor, S., Spapé, M., Biro, S., Hommel, B. (2013). From outcome prediction to action selection: Developmental change in the role of action-effect bindings. *Developmental Science, 16*, 801-814.

Chapter 8

Verschoor, S. A., Paulus, M., Spape, M., Biro, S. & Hommel, B. (Submitted). The developing cognitive substrate of sequential action control in 9- to 12-month-olds: Evidence for concurrent activation models.

Chapter 2: Means selection information overrides outcome selection information in goal attribution

Stephan Verschoor & Szilvia Biro

Abstract

It has been shown that, when observing an action, infants can rely on either outcome selection information (i.e. actions that expresses a choice between potential outcomes) or means selection information (i.e. actions that are causally efficient toward the outcome) in their goal attribution. However, no research has investigated the relationship between these two types of information when they are present simultaneously. In an experiment that addressed this question directly, we found that when outcome selection information could disambiguate the goal of the action (e.g., the action is directed towards one of two potential targets) but means selection information could not (i.e., the action is not efficiently adjusted to the situational constraints), 7- and 9-month-old infants did not attribute a goal to an observed action. This finding suggests that means selection information displays primacy over outcome selection information not only for adults but also for infants. The early presence of this bias sheds light on the nature of the notion of goal in action understanding.

Chapter 2

Introduction

To infer that an observed action is goal-directed (as opposed to random or accidental), i.e., that it is performed *in order to* achieve a particular outcome, one can make use of two types of information that can be directly observed in a given situation. One type of information consists of the observation of the way in which the actor adjusts the action to the situational constraints in order to achieve the end-state. Relying on this type of information implies the understanding of the relation between the means and the goal and the expectation that the actor will select the most efficient means action toward the goal. The other type of information consists of the observation that there are potential alternative outcomes and that the actor expresses a choice for a particular outcome via the action. We refer to these two types of information as "means selection" and "outcome selection" information, respectively.

While there are some situations in which only means selection or outcome selection information is present, in most situations both types are available. That is, it is often observed that an action is adjusted to situational constraints in order to achieve one of multiple potential outcomes. In this case both types of information contribute to a converging inference about the goal that the action is directed to. For example, stepping onto the bottom shelf in order to reach a bag of tortellini from the top shelf from among other types of pastas in a supermarket illustrates such a situation. Imagine, however, that you observe the same action, but that the tortellini is on the middle shelf (not on the top) and would thus be easily reachable without stepping onto the first shelf. In this case, the action expresses a choice between potential targets, but it does not seem to be adjusted efficiently to the situational constraints. The inference about the (goal-directedness and the) goal of the unexplained action (stepping on the shelf) will remain uncertain. Thus, when means selection information does not allow disambiguating the goal, even though the outcome selection information by itself could, no unambiguous goal-attribution can be made. Therefore, there seems to be a hierarchical relationship between these two types of information in our interpretation of goal-directed actions, namely means selection information is primary to outcome selection information. In other words, means selection information can override outcome selection information[1].

In this paper we investigate whether the relationship between the two types of information is similar in infants' interpretation of observed actions as in adults' interpretation, that is, whether means selection information also holds primacy over outcome selection information. Infants' ability to interpret observed actions as goal-directed has been investigated using two distinct experimental paradigms that have so far tested if and how infants can rely on one or the other type of information. However, until now no research has addressed the question of how infants interpret observed actions when the two types of information are present simultaneously.

Evidence for reliance on outcome selection information was found by showing that infants, after having watched an actor repeatedly acting upon (e.g. grasping) one of two objects, expect that the actor will continue to direct the action toward the same target after the spatial arrangement of the two objects has been altered (e.g., Woodward, 1998, 1999; Kiraly et al., 2003; Biro & Leslie, 2007). Furthermore, when outcome selection information is not provided in this experimental paradigm, that is, when the actor's action could not express a choice because there was only one object present, infants do not generate the expectation that the actor will act upon the same object after a novel object has been introduced (Baillargeon & Luo, 2005; Biro et al, 2012; Hernik & Southgate, 2012). Note that in these settings no (unambiguous) means selection information is available because the information about

---

[1] Note that we consider only directly observable information here. Additional, not directly perceivable or previous information about the actor (traits, habits, preferences, and beliefs) or the situation can alter the interpretation of means selection information and may allow disambiguating the goal (e.g., if you know that the actor has a habit of stepping onto the first shelf in the supermarket every time she grabs something).

situational adjustment of the action is non-sufficient (see Biro et al, 2012). Together, these findings suggest that outcome selection information is critical in infants' goal-attribution and that they rely on it from at least as early as 6 months.

The second experimental paradigm demonstrates infants' reliance on means selection information (Csibra et al., 1999, 2003; Sodian et al, 2004; Csibra, 2008). Infants are typically presented with a scenario in which there is only one potential outcome, which the actor repeatedly achieves by adjusting the action to the situational constraints (e.g., approaching a target by getting around an obstacle). When the situational constraints change, infants are found to predict that the actor will perform a novel and efficiently adjusted means action to achieve the original outcome (as opposed to repeating the familiar but no longer efficient means). If, however, the actor's action is not adjusted to the situation efficiently in the initial scenario, then infants do not generate a particular expectation about the new means action of the actor in a new situation. Thus, means selection information also has an important place in infants' early (also from around 6 months) goal-directed action interpretation.

In the present study we investigate how infants deal with situations which provide both types of information, means selection *and* outcome selection, to disambiguate the goal. Does means selection information have primacy over outcome selection information as in adult cognition? Or does outcome selection information dominate in infants' developing ability to attribute goals? The crucial situation, in which one can test this question, is when the two types of information lead to opposite conclusions, as in the tortellini example above. When an action is performed in the presence of multiple potential outcomes, but is not adjusted efficiently to the situational constraints, it remains ambiguous whether it was performed in order to achieve the outcome it led to. If outcome selection information has primacy, infants would interpret the action as directed to a particular outcome regardless of whether the means is inefficient toward this goal. If, on the other hand, means selection has primacy, then infants would not attribute a goal when the means is inefficient, even in the presence of outcome selection information.

To test this question, infants participated in three conditions. The Efficient Action condition involved familiarizing infants with a hand that first opens one of two transparent boxes that each contain a toy and then grasps the toy. In the non-efficient Gratuitous Action condition, the hand performs the same action, but the toys are in front of the boxes. Opening the box is therefore unnecessary and does not relate efficiently to the grasping[2]. In the non-efficient Superfluous Action condition, the boxes are not present, the hand grasps one of the toys after making some unnecessary movements in the air above and behind the toy (which actually mimic the opening of a box). The test events were identical for all conditions: two toys are present in swapped positions (no boxes) and the hand simply grasps either the same toy as before or the other toy. If outcome information has primacy over means selection information, then infants in all conditions should expect the hand to grasp the same toy as before. This expectation would be indicated by looking longer in the new toy test event than in the old toy test event in all three conditions. If means selection is the dominant source of information, then infants should attribute a particular goal and thus look longer in the new toy test event than in the old toy test event only in the Efficient Action condition, but not in the non-efficient conditions. Furthermore, the two types of non-efficient approaches allow us to investigate whether the inefficiency of the action can only prevent goal attribution if there is only one action involved (as in the Superfluous Action condition) or when the means can be considered as a part of a sequence, a sub action leading to an overarching goal (as in the Gratuitous Action condition). One could argue that inferring the goal by considering the inefficiency of a sub-action might be too demanding for young infants, and thus they might only rely in their interpretation on the second reaching action

---

[2] These two conditions were based on a previous study in which it was shown that 12-month-olds understand the causal relationship between opening a box and getting the object inside, and that they understand that opening the box is unnecessary if the toy is not inside but in front of the box (Woodward and Sommerville, 2000).

containing only outcome information. In the case of the presence of this transitional period, we would expect longer look in the new toy than in the old toy test event in the Efficient and the Superfluous Action conditions but not in the Gratuitous Action condition.

Method

Participants

Seventy-two 7-month-olds (mean age=31.10 weeks; SD=1.7 weeks; 36 girls and 36 boys) and seventy-one 9-month-olds (mean age=39.60 weeks; SD=1.08 weeks; 39 girls and 32 boys) were randomly assigned to one of three conditions: Efficient Action (23, 24), Gratuitous Action (24, 24) and Superfluous Action (24, 24, respectively). An additional eight 7-month-olds and fourteen 9-month-olds were excluded due to experimental error (3, 7) and fussiness (5, 7, respectively).

Stimuli

Infants were shown video recordings (Fig.1). In all conditions infants saw 6 familiarization events. In the first familiarization event of the Efficient condition, two toys (a bear and a ball) were sitting on a stage which had a curtained opening on the right side. In the second familiarization event, infants were shown that the two toys were sitting inside two transparent boxes with red lids. In the 3-6 familiarization events a hand reached through the opening, opened one of the boxes by removing the lid, and grasped the toy inside. The hand remained still and kept grasping the toy until the end of the event. The Gratuitous Action condition was identical to the Efficient Action condition except that the toys were not inside but in front of the boxes. Thus, after the hand removed the lid, it grasped the toy in front of the box. In the Superfluous Action condition the boxes were absent during all familiarization events. In the 3-6 familiarization events the hand mimicked the action of removing a lid of a box and then grasped the toy in front of the location where the action had taken place.

Following familiarization, the same pre-test event was presented in all conditions. The boxes were absent and the two toys were shown in switched positions. This pre-test event was followed by two test events. In the Old toy test event the hand reached through the opening and simply grasped the same toy as in the familiarization events. In the New toy test event the new toy was grasped instead.

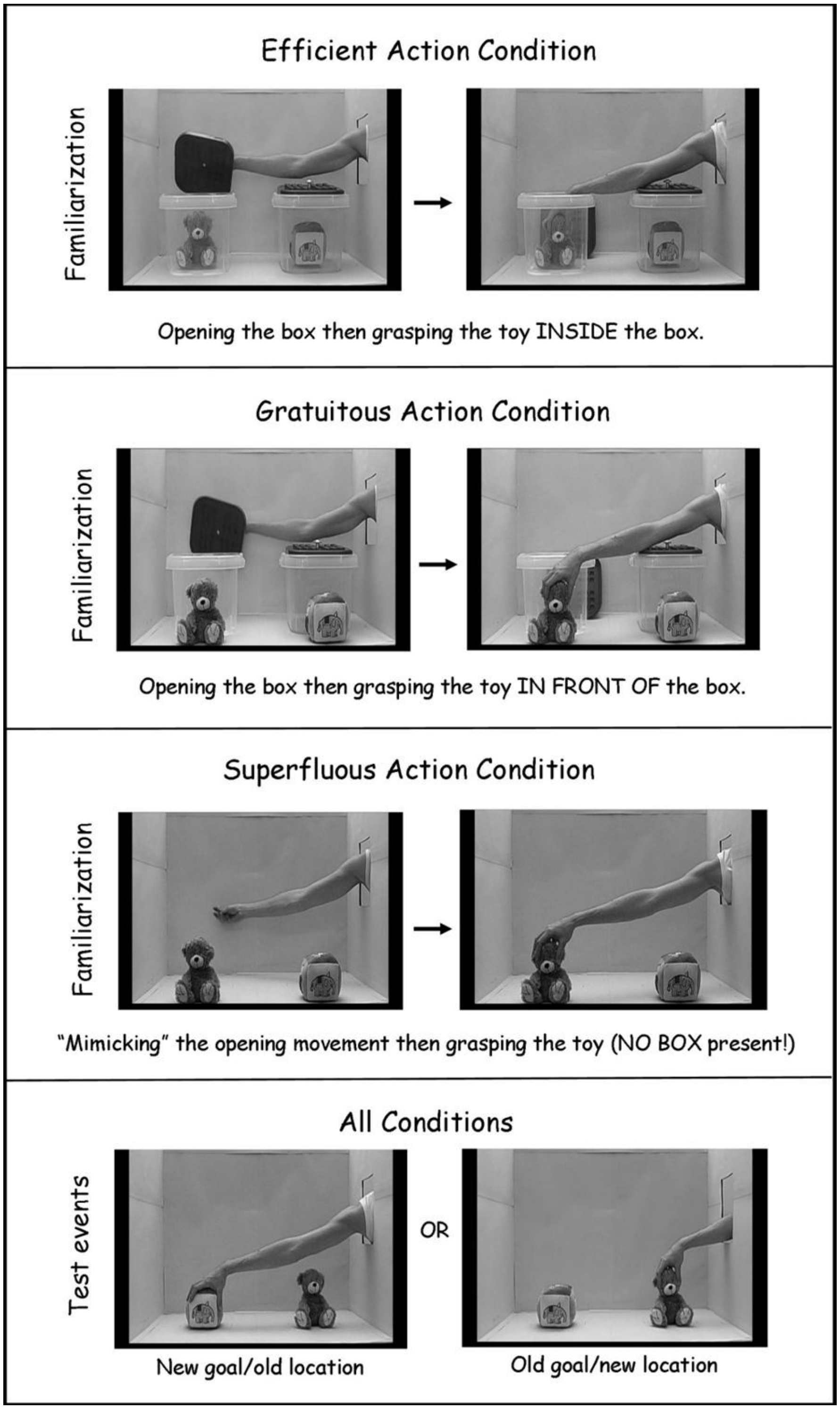

Figure 1. Illustration of the familiarization events 3-6 and the test events in the three conditions.

Apparatus and Procedure

During the experiment infants sat on their caregivers' lap in a curtained experimental booth. The stimuli were presented on a 31 inch widescreen monitor that was placed approximately 70 cm in front of infants' eye level. Two hidden cameras recorded the infants' faces and the stimuli. The on-line experimenter, who was blind to the condition, used specially built software (Schrama & Biro, 2005) to control stimulus presentation and to register infants' looking. At the start of trial the experimenter used an attention getter sound. When the infant was looking at the screen, the experimenter started the presentation of the event. If the infant watched the test event for less than 5 sec, then the trial was ignored and started again. (This was the minimum time required to ensure that the infants saw which toy was grasped in the test events.) If the infant looked away for more then 2 sec or if 120 sec had elapsed, the test event ended. Each infant saw both types of test events. The order of the presentation of the test events (New goal first/second), the type of object touched during familiarization (bear/ball), and the object's position (left/right side of the stage) were counterbalanced in all conditions.

The looking times of the test-phase and familiarization phase were re-coded off-line by a secondary coder. The inter-coder reliability for the test events was 96.86%.

Results

Looking times during familiarization were analysed first. A multivariate ANOVA revealed no significant difference between the three conditions and between the two age groups in the looking times for the familiarization and the pre-test events (Fig.2).

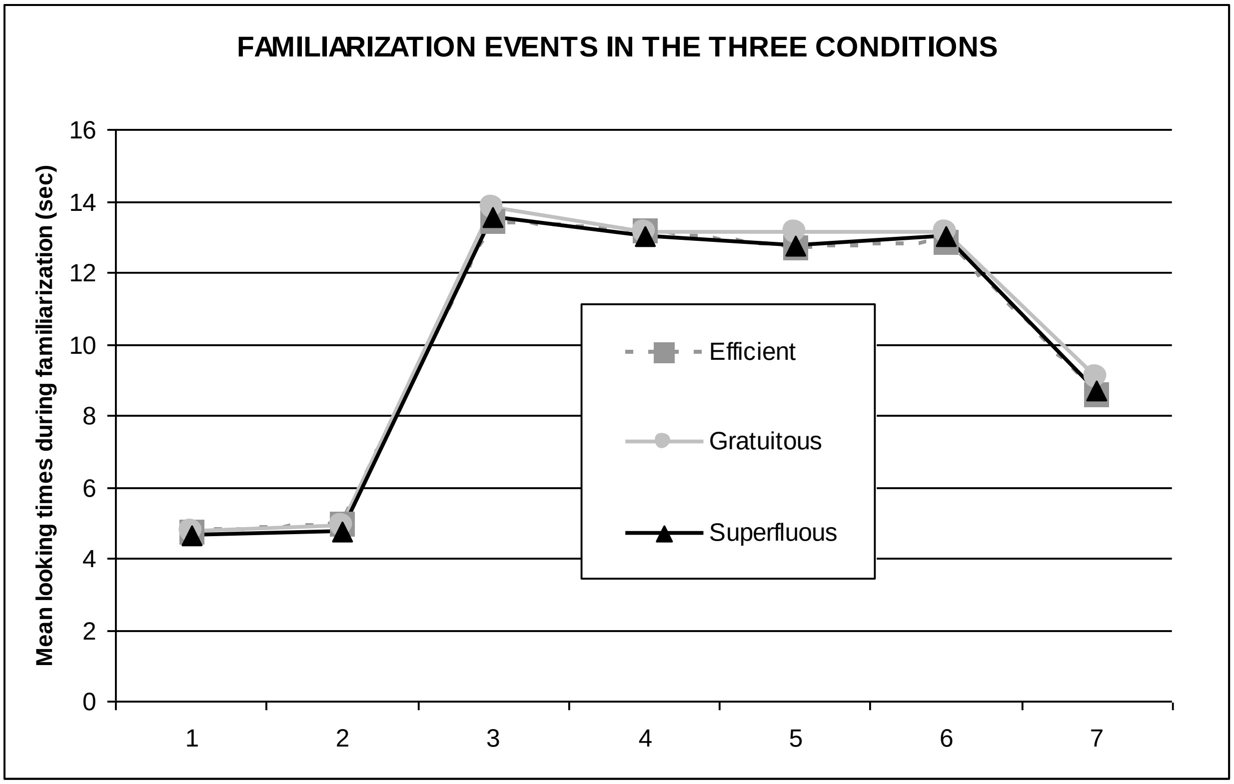

Figure 2. Mean looking times during familiarization for all conditions (s). The actual durations of the familiarization events were 5 seconds in events 1-2, 15 seconds in events 3-6 and the pre-test event lasted 10 seconds.

Second, the effects of order, gender, and toy position on the looking times in the test events were investigated. None of these factors had any main or interaction effect. These variables were therefore omitted from further analyses. A repeated-measures ANOVA was conducted on the looking times in the test events with condition, age-group and toy type (ball vs. bear) as between-subject variables. A significant main effect of condition was found [$F(2,131)=4.258$, $p=0.016$, effect size: 0.06]. Pair-wise comparisons indicated that infants overall looked longer in the Gratuitous Action than in the Superfluous Action condition ($p=0.005$). There was, however, no difference between the two non-efficient conditions and the Efficient condition. No main effects or interactions with age-group were found.

A strong interaction between test event type and condition was also found [$F(2,131)=7.139$, $p=0.001$, effect size=0.098]. Paired t-tests revealed that infants in the Efficient condition looked longer in the New toy test event than in the Old toy test event ($t(46)=-2.583$, $p=.013$, two-tailed). In the Gratuitous Action condition, the opposite pattern was found: infants looked longer in the Old toy test event than in the New toy test event ($t(47)=2.310$, $p=0.025$, two-tailed). No difference was found between the two types of test events in the Superfluous Action condition (Fig.3). A non-parametric Wilcoxon signed rank test confirmed these results (Efficient condition: $Z=-2.423$, $p=0.015$, Gratuitous Action condition: $Z=-2.595$, $p=0.009$).

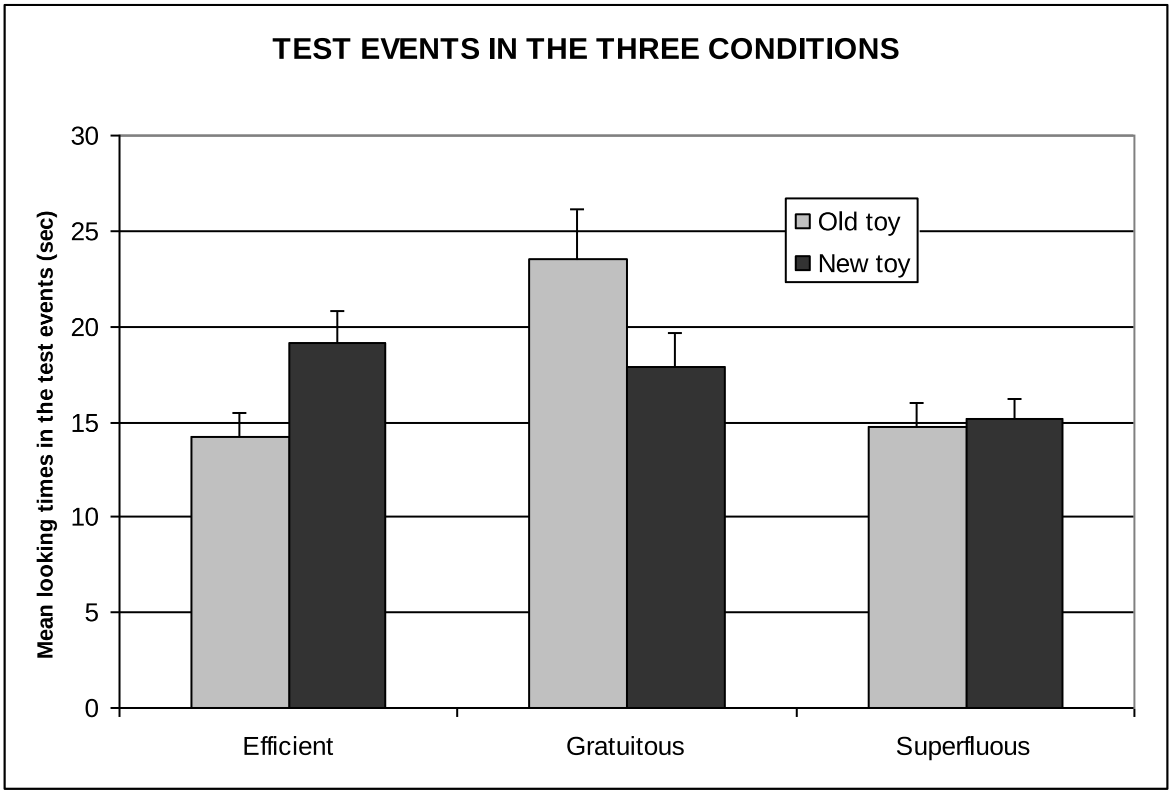

Figure 3. Mean looking times and standard errors (s) in the two test events by the three conditions. (* = $p < 0.05$)

Furthermore, a three-way interaction between test-type, condition and toy type was found [$F(1,131)=4.056$, $p=0.020$, effect size=0.058]. Follow-up tests revealed that only in the Gratuitous Action condition and only when the ball was the target, infants looked longer in the Old toy test event than in the New toy test event ($t(23)=3.130$, $p=0.005$, two-tailed), in the other conditions infants produced the same looking patterns in the two test events for both toys.

Discussion

We have investigated the nature of the relationship between outcome selection and means selection information in young infants' goal attribution. Infants were presented with situations in which both types of information were available to interpret and generate an expectation about the goal of an observed action. We found that 7- and 9-month-old infants only expected the actor to continue to direct the action toward the same target if both types of information were converging on the same goal-directed interpretation (Efficient action condition). In the other two situations, in which the outcome information could be used to disambiguate the goal but the means selection information did not justify this interpretation, infants did not generate a specific expectation about the actor's goal in a new situation (non-efficient action conditions). These findings together support the hypothesis that means selection information holds primacy over outcome selection information in infants’ action interpretation.

In neither of the two non-efficient conditions did the infants expect the actor to pursue the original outcome. This suggests that inefficiency of means action prevents goal attribution not only when there is a single action (with unnecessary detours, Superfluous condition), but also when the action can be considered to be part of an action sequence (Gratuitous condition). Our finding that infants as young as 7 months can evaluate the causal efficiency of a sub-action (opening the box) leading to an overarching goal (getting the toy) is quite remarkable as this ability had previously only been demonstrated at 10-12 months (Woodward & Sommerville, 2000; Sommerville & Woodward, 2005a). The early presence of this ability is also supported by previous evidence showing that by 5½ months infants can understand the physical constraints of containment events in goal attainment (Baillargeon, Graber, DeVos & Black, 1990), which is certainly necessary for the evaluation of the efficiency of the means in our containment event.

Recall that while the looking patterns in the non-efficient conditions were different from that in the Efficient condition, they were also different from each other. Looking equally long in the two test events in the Superfluous Action condition suggests that infants had indeed no specific expectation about which toy would be grasped. In the Gratuitous Action condition, infants looked significantly longer in the old toy than in the new toy test event. While this looking pattern also shows that infant did not expect the same goal to be achieved, the difference might reflect the fact that infants encoded the action in terms of its spatial properties (see also Woodward, 1998, Experiment 2). Infants might thus have expected to see the action to be performed on the same side of the stage as in the new toy event. One can speculate that the presence and the manipulation of the box in the Gratuitous Action condition highlighted the spatial position of the action, while the mimicked action in Superfluous Action condition did not tie the action as clearly to one side of the stage because no box was present to emphasize a particular location.

A further caveat that needs to be considered is whether infants in the Gratuitous Action condition might have evaluated the action of opening the empty box as being efficient action - despite our intention for it to look inefficient - by inferring that the goal of the action was to put a toy *into* the box[3]. Since the boxes were no longer present in the test events, this interpretation was not tested in the test events and this possibility can thus neither be verified nor rejected. However, we argue that this interpretation is quite unlikely, because it requires the inference of a hypothetical state of affairs since this goal has not been seen achieved. Inferences about hypothetical scenarios are above the level of the representational ability

---

[3] Adult viewers sometimes came up with this interpretation of the familiarization events.

of infants at the age tested in the current study (Csibra et al., 2003; Bellagamba & Tomasello, 1999). Furthermore, in the familiarization events, after opening the box the hand kept grasping the toy for 5 sec until the end of the event, which would be unnaturally long if the goal had been putting the toy into the box. Note, however, that such an interpretation of the familiarization phase of the Gratuitous Action condition would have involved relying on the means selection information rather than on the outcome selection information. Hence, it would not change our main argument about the relation between the two types of information.

Our findings provide a crucial piece of supporting evidence for a proposed theory about the mechanism underlying infants' understanding of goal-directed actions. This theory states that infants are equipped with an abstract interpretational system, the teleological stance, which represents goal-directed actions by relating three elements: the observed action, the end state of the action, and the physical context in which the action takes place (Gergely & Csibra, 2003). The central claim of this theory - that infants *only* attribute a goal to an action if the action can be evaluated as an efficient action toward the end-state in the given situation - has so far only been supported by relatively indirect evidence. The experimental paradigm that had been used to investigate this theory tested infants' predictions about the means, and not the goal, of the action. However, the finding that infants do not generate a specific expectation about the new means if the initial action is non-efficient only suggests, but does not prove, that infants do not in the meantime attribute a goal to the initial action. Infants may have given up on predicting the means, but could in principle have relied on the end-state of the action to infer it as the goal. Our current setting enabled us to directly test the expectations about the goal of the action by providing two possible end-states. Our finding thus confirms that if the action, the end-state, and the situational constraints do not form a relation that satisfies the principle of efficient action, then infants do not commit themselves to a specific goal.

Another related point that needs to be addressed is how our design and research questions differ from those of Woodward and Sommerville (2000). In that study 12-month-olds were shown the same familiarization events as in the Efficient and the Gratuitous Action condition of the present study. However, in the test events, only the two means actions were shown and not the goal attainment: the boxes were still present, the toys were either inside or in front of the boxes, and the hand only touched one of the boxes without grasping the toy. Therefore, the finding that infants did not have specific expectations about the means in their non-efficient condition did not necessarily prove that they did not interpret the grasping of one of the toys in the familiarization phase as goal-directed. While both the current and Woodward and Sommerville's experiments provided simultaneous presence of outcome selection information and means selection information, only the present study clarified that means selection information overrides outcome selection information in specifying the goal of the observed action.

Our findings reveal that means selection information has priority over outcome selection information from a very young age. It is an interesting question whether there exists any period during the development of action interpretation in which infants would preferentially rely on outcome selection information if means selection information is also available. On the one hand, processing means selection information is more demanding in terms of computational resources and background knowledge, which would suggest that at the very beginning infants might primarily rely on encoding the end-state of an action in goal attribution. On the other hand, from birth infants perceive a flow of action sequences in an ever-changing rich environment. To interpret every perceptually salient end-state as a goal, even if that interpretation is not justified by the means in the given context, would be counterproductive since it could lead to misinterpretation and false action predictions, and would certainly not allow infants to start to parse the flow of actions into meaningful and hierarchically related goal-directed action units (Baldwin, Baird, Saylor & Clark, 2001).

The demonstration of the primacy of means selection information does not imply that outcome selection information is unimportant in action understanding. On the contrary, it is very useful information when it is the only observable information one can rely on and it is very likely crucial for understanding and attributing preferences. In fact, there is an on-going debate about whether the experimental paradigm that investigates sensitivity to outcome selection information alone (Woodward, 1998) is a test for preference attribution rather than for goal attribution or both (Luo & Baillargeon, 2005; Hernik & Southgate, 2012; Jacob, 2012). In any case, recent studies have demonstrated that 9- and 12-month-old infants can transfer their goal attribution from a situation that contains only means selection information to another situation in which only outcome selection information is available (Biro et al, 2012; Hernik & Southgate, 2012). These findings suggest that these two types of information feed into a unitary goal concept.

In sum, our current study shows that infants' early understanding of goal-directed actions is similar to that of adults as far as the relationship between means selection and outcome selection information is concerned. The early presence of the preferential bias for means selection information suggests that this bias is not an acquired, but rather a core property of the cognitive mechanisms of goal attribution and thus it sheds light on the nature of the notion of goal.

Chapter 3: Evidence for a unitary goal concept in 12-month-old infants

Szilvia Biro, Stephan Verschoor, Lot Coenen

Chapter 3

Chapter 3

Abstract

We investigated whether infants can transfer their goal attribution between situations that contain different types of information about the goal. We found that 12-month-olds who had attributed a goal based on the causal efficacy of a means-end action, generated expectations about the actor's action in another scenario in which the actor could choose between alternative outcomes. This finding suggests that, by 12 months, infants possess a unitary concept of goal.

Chapter 3

Chapter 3

Introduction

Goal-directed actions never take place in an empty space, they are embedded in a particular situation. To interpret an observed action as goal-directed, that is, to infer that its outcome is not random or accidental but carries explanatory power about the action, one can make use of two types of information in a given situation. One type of information consists of the observation that, to achieve the goal, the actor adjusts the action to the situational constraints. We refer to this type of information as "means selection". The other type of information is the availability of alternative outcomes and the expression of a preference for a particular outcome by the actor's action ("outcome selection").

Infants' ability to interpret actions as goal-directed has been investigated using two distinct experimental paradigms that have separately tested if and how infants can rely on one or the other type of information. The first paradigm, introduced by Amanda Woodward (1998), assesses infants' sensitivity to outcome selection. Typically, infants are habituated to an event in which two distinct toys sit on a stage and an actor repeatedly acts upon, for example grasps, one of the toys. Thus, infants observe that the actor consistently chooses one of two potential targets. In the test events the locations of the toys are swapped and the actor acts upon either the same toy or the other toy. Infants are found to look longer in the other toy than in the same toy test event. This looking pattern (combined with appropriate controls that excluded simple association) is interpreted as evidence for infants' ability to encode the action as directed to a particular outcome and to expect that the actor will continue to act to achieve the same outcome.

Further evidence that outcome selection is critical for infants' expectations in this paradigm comes from a study showing that if outcome selection information is not provided in the habituation/familiarization phase, infants do not seem to interpret the action as goal-directed (Luo & Baillargeon, 2005). In this variation, only one toy is present during the familiarization phase when the actor repeatedly touches the toy. A second, novel toy is only introduced before the two test events in which the actor acts upon either the same toy as before or the novel toy. Infants look equally long in the two tests events, showing that they do not develop an expectation about which toy the actor would act upon. Thus, when outcome selection information is not provided, infants do not make goal-directed action predictions[4].

The second paradigm investigates infants' reliance on means selection information (e.g., Gergely et al., 1995; Csibra et al., 1999, 2003; Csibra, 2008; Wagner & Carey, 2005; Sodian et al., 2004). Infants are typically habituated to a scenario in which there is only one potential outcome, but which the actor achieves by adjusting the action to the situational constraints (e.g., approaching a target by getting around an obstacle). In the test events, the situational constraints change (e.g. the obstacle is removed) and the actor either adjusts his/her behavior to the new situation or performs the same means action, which, however, is no longer efficient with respect to the outcome. Infants' looking time pattern in the test events suggests that they interpret the action as goal-directed and expect the actor to perform the novel and efficiently adjusted means action to achieve the goal. However, if no means selection information is included, either because the actor’s action does not appear to be efficiently adjusted to the given situation constraints during the habituation phase (e.g. the actor takes a detour even though a more direct and shorter route is available) or because the action does not show any adjustment due to a lack of situational constraints, infants do not develop specific expectations for the actor's novel means to achieve the outcome (e.g. Csibra et al., 2003; Biro et al., 2007). These findings suggest that, in this paradigm, infants only attribute goals to actions when unambiguous information about means selection can be obtained.

The differences between the two paradigms in terms of the type of information they provide to interpret and predict actions, raise a question about the unitary nature of the goal concept that infants

---

[4] The interpretation of the results on the one-target version (Luo and Baillargeon, 2005) is further discussed in Henrik and Southgate's paper (same issue), but from a different angle.

apply in these situations. It is possible that infants appeal to one goal-like notion (e.g., 'preference') in situations that provide outcome selection information, and to another one (e.g., 'planning') when they receive evidence on means selection. Alternatively, a unitary concept of goal explains the results of both kinds of studies, which takes input from either type of information. In the latter case, one would expect transfer of goal attribution from situations with one type of information to situations in which the other type of information is available.

Such transfer between situations with the two different types of selection information often occurs in everyday life when we make predictions about others' actions. Consider, for example, the question of a disappointed wife to her husband: "You went to the Moon and back in order to marry me and now you leave me for your young secretary?" This question clearly expresses the expectation that the husband, who, in the past, apparently showed that he could efficiently overcome large obstacles to achieve his desired goal (means selection information was used to justify the goal), would continue to desire the same goal even if other potential desirable objects appeared in a future situation (when outcome selection information became available to make a goal-directed action prediction).

Our aim in the present study was to combine the two paradigms in order to test whether infants can transfer goal attribution from a situation in which only means selection information is present to a situation in which outcome selection information is available.

In particular, we investigated whether infants are able to attribute a goal in the one-toy variation of the Woodward paradigm (for which it was previously found that they could not, Luo and Baillargeon, 2005) if means selection information were provided during the familiarization phase. In other words, does observing an actor efficiently adjusting the action to situational constraints lead infants to expect that the actor will continue to approach the same goal after another potential outcome is presented in the scene?

Infants participated in three conditions. The Efficient action condition involved familiarizing infants with a hand that opened a transparent box and grasped the toy inside. This condition was contrasted with two control conditions. In the Non-efficient action condition the hand performed the same action, but the toy was in front of the box. Opening the box therefore was unnecessary and did not relate efficiently to the grasping[5]. In the second control condition, the Simple action condition, the box was not present, and the hand simply grasped the toy. Grasping an object does not qualify as an inefficient action per se, at least for an adult viewer. However, a simple grasping action does not provide unambiguous information on means selection because it does not show adjustments to situational constraints. The test events were identical for all conditions: two toys were present and the hand grasped either the new or the old toy. If infants can transfer goal attribution from means selection to outcome selection, they should only look longer in the new toy than in the old toy test event in the Efficient action condition.

Method

Participants

---

[5] These two conditions were based on a previous study in which it was shown that 12-month-olds understand the causal relationship between opening a box and getting the object inside, and that opening the box is unnecessary if the toy is not inside but in front of the box (Woodward and Sommerville, 2000). However, note that our study differs from the one of Woodward and Sommerville's both in its design and in its research question. In their study, two different boxes were used with two toys during the habituation (instead of a single box with a single toy as in the current study). Thus, it could be argued that both types of selection information (means and outcome) were present in their study. Furthermore, only the means action (touching the box) was shown in the test events and not the goal attainment (grasping the toy, as in the current study). The question they asked was: are babies able to infer the necessary means that previously led to the goal in a changed situation?

Fifty-two full-term 12-month-olds (mean age = 52.44 weeks; SD=1.4 weeks; 24 girls and 28 boys) participated in the experiment. An additional eight infants were excluded due to experimental error (3) or fussiness (5). The infants were recruited through direct mail. Infants received toys and certificates as gifts and parents could opt for compensation for travel expenses. The infants were randomly assigned to three conditions: Simple action (16), Efficient action (18) and Non-efficient action (18).

Stimuli

Infants were shown video recordings. In all conditions, infants saw 6 familiarization events. In the Simple action condition, the first and the second familiarization events (each 5 sec) showed a puppet stage with a curtained opening on the right side. A toy (bear or ball) was sitting at one side of the stage (left or right). In the 3-6 familiarization events a hand reached through the opening and grasped the toy (Figure 1). The hand remained still and continued to grasp the toy until the end of the event (15 seconds each). The Efficient action and Non-efficient action conditions were identical to the Simple action condition except that the toy in the second familiarization event was, respectively, inside or in front of a transparent box. During the 3-6 familiarization events the hand reached through the curtain, opened the lid of the box and then grasped the toy inside/in front of the box. These events also lasted 15 seconds, but the duration of the phase in which the hand was in motion was 5 seconds longer than in the Simple action familiarization events.

Following familiarization, a pre-test event was presented. In all conditions the familiar toy was shown on the opposite side of the stage while a novel toy was sitting in the familiar place (10 sec). This pre-test event was followed by two test events. In the Old toy test event the same toy was grasped as in the familiarization events, whereas the novel toy was grasped in the New toy test.

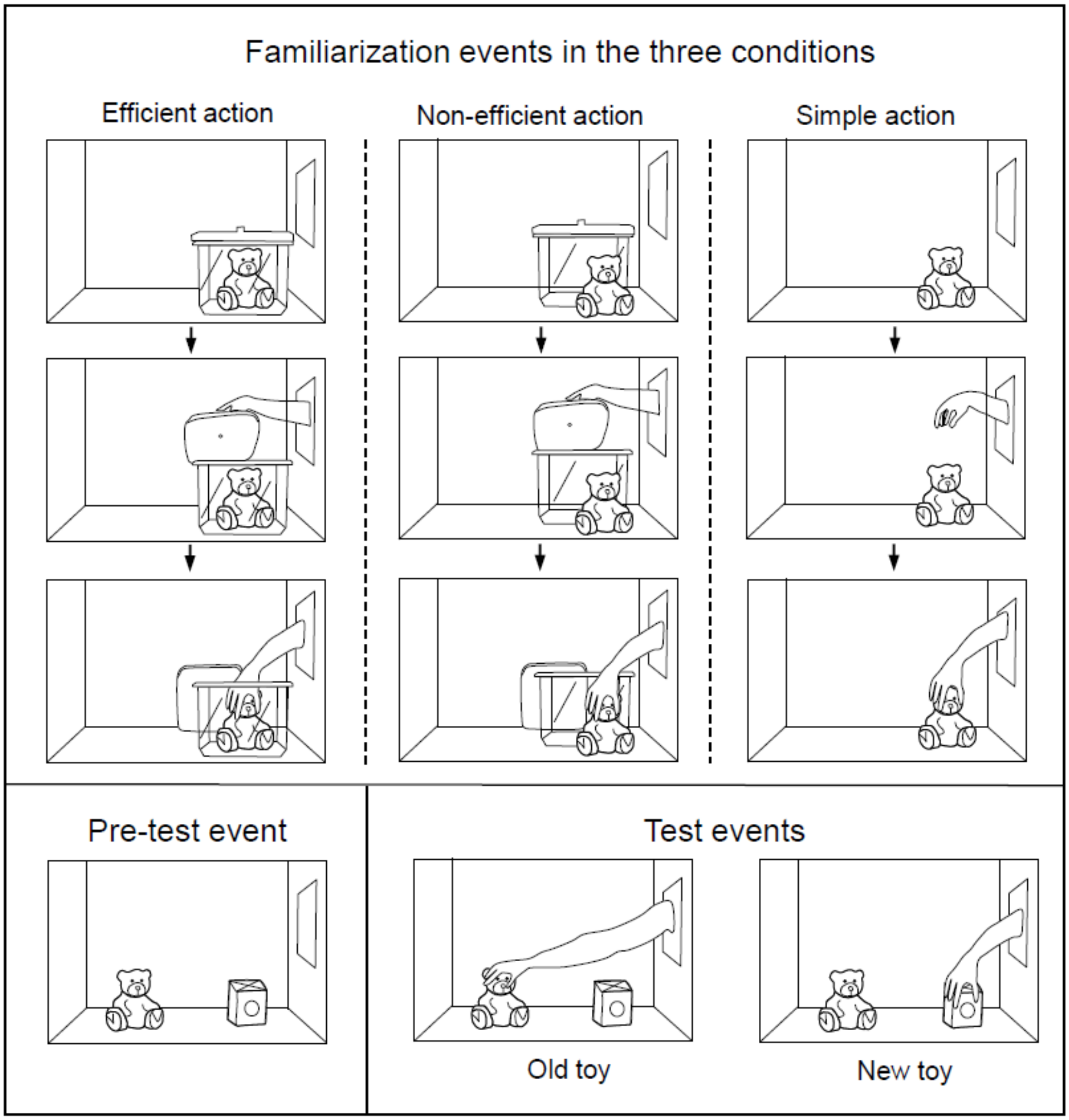

Figure 1: Illustration of the familiarization events 3-6, pre-test and test events in the three conditions.

Apparatus and Procedure

During the experiment the infant sat on their caregivers' lap in a curtained experimental booth. The stimuli were presented on a 31 inch widescreen monitor that was placed approximately 70 cm from the infant’s eye level. A hidden camera located above the monitor recorded the infant’s face. Another camera behind the infant recorded the stimulus presentation. The cameras fed a mixer that produced a split-image movie of the infant and stimuli. The on-line experimenter, who was blind to the condition, used specially built software (Schrama & Biro, 2005) to control stimulus presentation and to register the infant’s looking. At the start of the session the experimenter used an attention getter (a doorbell type of sound). When the infant was looking at the screen the experimenter started the presentation of the familiarization event. The attention getter was used before each trial. The experimenter measured the infant's looking time in the test events on-line by pressing a key when the infant was looking at the monitor. If the infant did not watch the test event for at least 5 s, then the trial was ignored and started again. (This was the minimum time required to ensure that the infants saw which toy was grasped in the test events.) The test event ended if the infant looked away for more than 2 s or if 120 s had elapsed.

Each infant saw six familiarization events, one pre-test event and both types of test events. The order of the presentation of the test events (New toy first/second), the type of object shown during familiarization (bear/ball), and the position of the object (left/right side of the stage) were counterbalanced in all conditions.

The looking times during the familiarization and test events were also measured by an off-line coder. The inter-coder reliability for the test events was 97.05%.

Results

Looking times during familiarization were first compared for the three conditions. A one way ANOVA revealed that there was a significant difference between the average looking times in the 3-6 familiarization events [$F(2,51) = 7.892$, $p = 0.001$]. Post-hoc comparison (LSD) showed that this effect was due to the fact that, in the Simple action condition, infants looked less at the familiarization events than in the Efficient ($p < 0.0001$) or Non-efficient action conditions ($p = 0.032$), while these latter two conditions did not differ from each other. There was no difference between the conditions in looking times for the first two familiarization and the pre-test events. (See Figure 2 for mean looking times.)

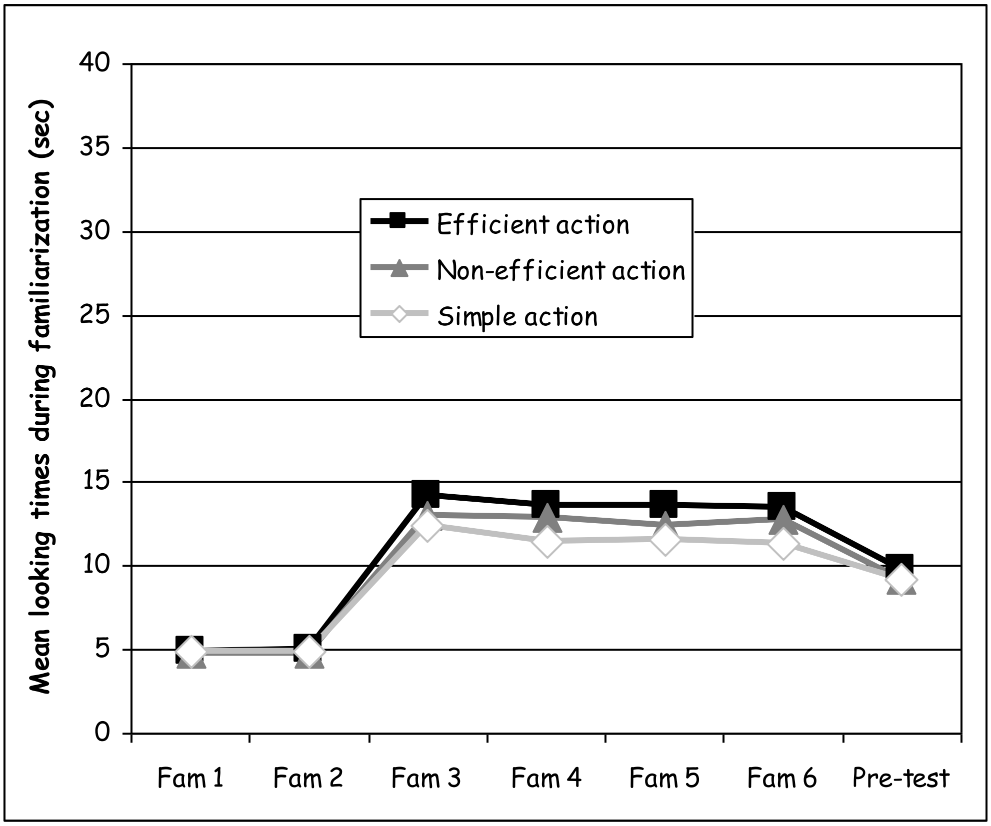

Figure 2: Mean looking times during familiarization (s). The actual durations of the familiarization events were 5 seconds in events 1-2, 15 seconds in events 3-6 and the pre-test event lasted 10 seconds.

Second, the effects of the order, gender, toy type and toy position on the looking times in the test events were investigated. None of these factors had any main or interaction effect. These variables therefore were omitted from the main analyses. A repeated measures ANOVA was conducted on the looking times in the test events with condition as a between-subject variable. A significant main effect of test event type was found, indicating that overall infants looked longer in the New toy test event than in the Old toy test event [$F(1,49) = 5.997$, $p = 0.018$, effect size = 10.9%]. An interaction between test event type and condition was also found [$F(2,49) = 4.305$, $p = 0.019$, effect size = 14.9%]. Paired t-tests revealed that infants in the Efficient condition looked longer in the New toy test event than in the Old toy test event ($t(17) = 3.059$, $p = 0.007$, two-tailed). There was, however, no difference between the two test events in the Simple action and in the Non-efficient action conditions (see Figure 3). A non-parametric Wilcoxon signed rank test confirmed these results (Efficient condition: $Z = 2.461$, $p = 0.014$).

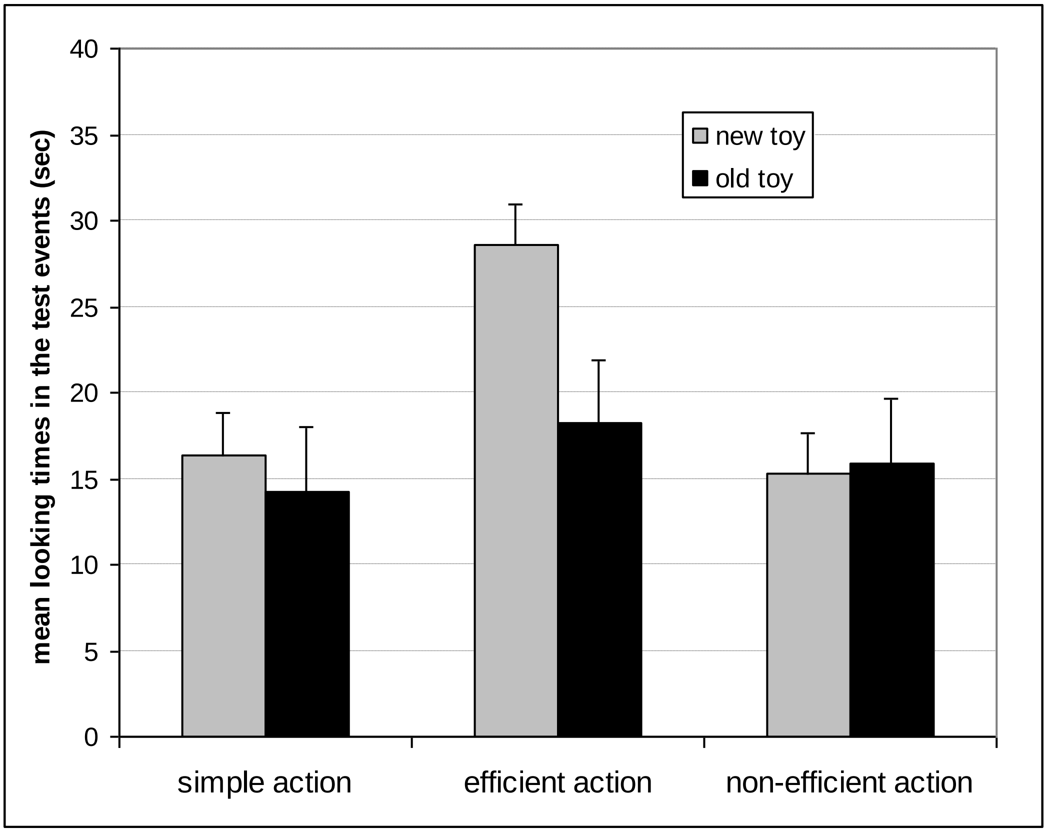

Figure 3: Mean looking times and standard errors (sec) in the two test events by the three conditions. (* = $p < 0.05$)

Discussion

In the present study we investigated whether observing an actor who engages in an efficient means action towards a goal, leads infants to expect that the actor will continue to approach the same goal even when another potential outcome is present. We found that in the Efficient means action condition infants looked longer in the new toy test event than in the old toy test event, that is, they expected the action to be directed towards the same target in the new situation. In the Non-efficient and the Simple action conditions, however, infants generated no expectation about the possible target of the action, as their looking times for the new and old toy test events were not different. These findings suggest that infants can transfer their goal-directed interpretation and prediction of action from a situation containing means selection information to a situation with outcome selection information.

The lack of looking time difference between the two test events in the Simple action control condition serves not only as a replication, but also as an extension of the previously mentioned finding in the one-target variation of the Woodward paradigm (Luo & Baillargeon, 2005). In that study the actor was an inanimate self-propelled box that approached the toy in a straight line movement. Thus, our null result shows that goal-directed interpretation of the action in the one-object familiarization setting is not only disrupted in the case of inanimate actions, but also for human hand action[6]. In comparison to the Efficient action condition, the null result in the Simple action also suggests that a simple grasping action does not contain sufficiently unambiguous means selection information to compensate for the lack of outcome selection information.

The null result in the Non-efficient control condition indicates that goal-directed interpretation is also prevented if the means action is clearly not efficient in the given situation toward the outcome. Furthermore, this null result is also important because it excludes alternative, lower-level explanations for the difference in looking times between the Efficient and the Simple action conditions. For example, saliency or increased attention due to the longer and more complex actions during familiarization in the Efficient condition compared to the Simple action could in principle have caused the different looking

---

[6] Lou and Baillargeon (2007) reported an unpublished study (Song et al., 2006) finding a similar null result in a one-object setting in which a hand grasped and lifted the object. In addition, Luo and Baillargeon (2007) and Luo and Johnson (2009) showed that the absence of “outcome selection” information from the perspective of a human actor also prevents infants from generating expectations about a goal-directed action of the actor. In their studies two toys were presented, but the perceptual access of the actor was blocked for one of the toys.

patterns in these two conditions for the test events. However, none of these factors can account for the difference between the Efficient and Non-efficient conditions.

Overall, our findings demonstrate that infants are able to transfer goal attribution based on means selection information to a situation in which outcome selection information is available. Furthermore, the present study also adds to previous findings that demonstrated the critical role of infants' sensitivity to the efficiency of the means in goal-attribution.

The present study raises two questions about the relationship between the two types of information that infants utilize in their interpretation of actions. One question concerns the direction of transfer in goal-directed interpretation. Our study demonstrated transfer from situations with means selection to situations with outcome selection information. Would the transfer also work the other way around? In other words, would providing infants with outcome selection information lead them to expect that in a changed situation the actor will engage in an efficiently adjusted action to achieve the original goal?

The second issue concerns the primacy of one of the two types of information in goal attribution. Which type of information would infants rely on if both types were available in one situation? Two studies have varied the efficiency of the means in the standard two-object setting of the Woodward paradigm. Woodward and Sommerville (2000) showed that 12-month-olds only develop an expectation about a specific means action if this means was previously shown to be causally (and efficiently) related to the goal of the action. However, this study did not test the goal attribution directly. Biro and Verschoor (2007), however, found that 7-12 months old infants only expect the same goal to be achieved if the means was efficiently related to the outcome. These findings suggest that means selection information may overwrite the outcome selection contrast, that is, infants only consider the outcome of the action in their goal-directed action representation if the means that led to the outcome was efficient.

For the two-target setting it has also been shown that the outcome selection information is not exploited if the action is performed by an inanimate object (Woodward, 1998) or if the action is unfamiliar (Woodward, 1999; Guajardo & Woodward, 2004). However, when these types of actions were enriched with certain movement characteristics, infants did develop particular expectations of goal-directed actions of the actor (Kiraly et al., 2003; Biro & Leslie, 2007; Johnson, Shimizu & Ok 2007a). One can speculate that these movement cues (variation of action, self-alignment, salient effect) served as hints for means selection information and could thus replace (similarly to our present study), or at least facilitate, the use of outcome information. Furthermore, in a one-target situation with a non-human actor, infants only interpreted the approached target as the goal of the action, and expected the non-human actor to approach the same target when another target was also available, if the action was efficiently related to the target (Hernik & Southgate, 2012). This study independently corroborates and extends our finding as it indicates that transfer of goal attribution between situations providing two different types of information occurs even in the case of non-human actors.

Finally, our finding that 12-month-old infants can transfer goal attribution from a situation in which one type of information is available to a situation in which the other type of information is present, suggests that they may possess a unitary goal concept that can take input from different types of evidence to generate expectations for goal-directed actions. Such a unitary notion of goal may allow infants to recognize the "common denominator" in different aspects of human behavior, such as planning or the expression of preferences. Both planning (based on selection between means) and preference (based on selection between outcomes) imply that the actor is making a choice regarding her/his actions. Recent theoretical proposals have given a prominent role to the assumption of "choice making power" in the understanding of agency (Leslie, 2010; Gergely, 2009). By showing that infants can connect these two aspects of choice making, both our current study and that of Hernik and Southgate (2012), support the idea that infants indeed rely on this assumption in order to identify agents and to make sense of their behavior.

Chapter 4: Outcome potential influences twelve-month-olds’ interpretation of a novel action as goal-directed

Szilvia Biro, Stephan Verschoor, Esther Coalter and Alan M. Leslie

Chapter 4

Chapter 4

Abstract

Learning about a novel, goal-directed action is a complex process. It requires evaluating the causal potential of the action and linking the action to its outcome for later use in new situations to predict the action or to anticipate its outcome. We investigated the hypothesis that linking a novel action to a *salient change in the environment* is critical for infants to assign a goal to the novel action. We report a study in which we show that 12-month-old infants, who were provided with prior experience with a novel action accompanied with a salient visible outcome in one context, can interpret the same action as goal-directed even in the absence of the outcome in another context. Our control condition shows that prior experience with the action, but without the salient effect, does not lead to goal-directed interpretation of the novel action. We also found that, for the case of 9-month-olds infants, prior experience with the outcome producing potential of the novel action does not facilitate a goal-directed interpretation of the action. However, this failure was possibly due to difficulties with generalizing the learnt association to another context rather than with linking the action to its outcome.

Introduction

As adults, we habitually see actions of social agents as bent toward particular ends. However, learning about a novel, goal-directed action is not a straightforward process, it involves multiple steps. When one observes a novel action for the first time, one has to be able to identify the outcome of the observed action and figure out which aspects of the action are essential for achieving the outcome, that is, one has to evaluate the causal potential of the action. A second necessary step is to store the link between the observed novel action and its outcome for future use. Finally, one has to be able to retrieve the stored link and to use it in an entirely new context. For example, observation of the novel action in a new situation can lead to the anticipation of the outcome without having witnessed the attainment of the associated outcome. This last step of appreciating that the novel action has the potential to produce a particular outcome, enables us to engage in fast on-line processes of intertwined goal anticipation and action prediction[7].

A number of mechanisms have been suggested for the above steps in the process of learning about a novel goal-directed action. Teleological reasoning (Gergely & Csibra, 2003) and simulation procedures (e.g. Meltzoff, 2002; Tomasello, 1999) have been suggested to solve the inferential problem of action and goal selection, and a bi-directional action-effect association mechanism (Hommel et al., 2001) has been proposed to take care of the linking and retrieval of stored means-end relations. While there is debate about the relations between these mechanisms and the role they play in the emergence and

---

[7] Note that by the term "outcome" we mean the actual (causal) effect of the action, while we use the term "goal" to refer to the mental representation of this outcome. An action is thus considered goal-directed if it can be seen as performed to achieve a particular outcome, in other words, if it is a means to an end.

development of goal-directed action interpretation in infancy (e.g., Woodward, 2009; Biro & Leslie, 2007; Biro et al., 2011; Shimizu & Johnson, 2004; Johnson, 2000), most of these mechanisms seem to share an assumption about an important precondition in infancy for interpreting novel actions as goal-directed. For infants to be able to successfully identify the outcome, and to link it to the action, the outcome of the novel action needs to involve a *salient and easily detectable change* in the environment (e.g. Király et al., 2003; Elsner, 2007, Biro & Leslie, 2007; Meltzoff, 1988; Verschoor et al., 2010).

The assumption of the necessity of a salient outcome is inherent in teleological reasoning (Gergely & Csibra, 2003), which states that an action is judged as a well-formed goal-directed action if it can be justified as an efficient action towards the outcome. Thus, when one observes a novel action, the goal of the action can be inferred by considering what change of state would be efficiently brought about by this action in the given situation. However, young infants have only limited background resources – such as knowledge about physical constraints - to make such inferences when the outcome is not immediate, not directly visible, or only one of many co-occurring outcomes (e.g., Csibra et al., 2003). The presence of a salient effect can therefore considerably ease the evaluation of the causal potential of an observed novel action. The role of salient changes is also central in the theoretical account proposed to explain the development of action planning and perception (Elsner & Hommel, 2001) on the basis of William James' ideomotor principle. This theory states that actions are inherently represented and linked to each other by their distal effects. Furthermore, theories that credit infants with domain specific modular systems that are sensitive to behavioral cues have also emphasized the role of salient outcomes caused by the means in infants' attribution of goal-directedness (e.g. Leslie, 1994; Gergely et al., 1995). In particular, Biro and Leslie recently proposed a cue-based bootstrapping model (2007) which claims that infants’ innate sensitivity to behavioral cues (such as salient outcomes, self-propelledness, variations of the action) is coupled with a learning mechanism that can link behavioral cues to other surface features such as the appearance or type of actions. When such a link has been established, infants can anticipate goal-directed actions from actors identified by surface features without collecting direct evidence for behavioral cues.

While the influence of observed salient action effects on infants' own exploratory or imitative behavior is well-documented (e.g. Elsner, 2007; Elsner and Aschersleben, 2003; Hauf et al., 2004, Rovee-Collier, 1987; Verschoor et al., 2010), there is much less evidence of their impact on infants' understanding of and prediction for others' observed novel actions (see discussion below). The aim of the current paper is to investigate the role of a salient outcome in the process of interpreting a novel action as goal-directed. Thus, we asked whether infants can appreciate the link between a novel action and its salient outcome and whether they can make use of the learnt link in a new context by interpreting the observed novel action as directed to a particular goal even in the absence of the salient outcome. In other words, we tested whether infants can assign a goal to an unfamiliar action that in a different context they had previously associated with a salient visible effect.

There is evidence that infants can already interpret a *familiar* hand action such as grasping as goal-directed as early as six months of age. In a study using a visual habituation method, infants watched a hand repeatedly grasping one of two toys on a stage (Woodward, 1998). After habituation, the positions of the toys were swapped and the hand grasped either the new toy in the old location or the old toy in the new location. Infants looked longer at the new toy test event indicating that they expected the hand to grasp the same toy. However, when 6 and 9-month-old infants were habituated in a similar setting with an unfamiliar novel action in which the back of a hand simply touched one of the target toys, the infants looked equally long in the two test events, which suggests that they did not specifically expect the hand would touch the same toy again (Woodward, 1999).

It has been suggested, however, (Király et al., 2003, Biro & Leslie, 2007) that infants' predictions in the grasping hand condition might not have been based on the grasping action per se, but rather on the strong association of the potential salient outcome that a grasping action can produce (such as the picking up of an object), a means-end relation with which infants are very familiar (Leslie, 1982, 1984). The crucial role of a salient action effect in interpreting a novel action as directed towards a particular goal-state has been demonstrated by replicating the "back of the hand touch" experiment in which the back of the hand did not only touch, but also pushed the target toy to a new location in both the habituation and test phases (Király et al., 2003; Jovanovic et al., 2007). In these experiments, infants from 6 months seemed to interpret the action as goal-directed: they expected the hand to touch and push the same object. A similar looking pattern was found when other types of unfamiliar actions with salient outcomes were used in the same experimental paradigm (Hofer et al., 2005; Biro & Leslie, 2007), for example when a wooden rod touched and - with the help of a Velcro piece - picked up the target toy.

Woodward and colleagues (Heineman-Pieper &Woodward, 2003; Cannon & Woodward, 2010) have criticized these studies, however, by arguing that moving a new object can explain longer looking time to the new toy test event rather than a violation of expectation regarding the goal of the action. However, in the Biro and Leslie study (2007), the force of this alternative explanation is considerably weaker. In that study, 9- and 12-month-old infants looked longer in the new toy than in the old toy test events, even when there was movement only during the familiarization but not in the test events, in which the toy was touched only.[8] This finding also indicates that 9- and 12-month-old infants were able to generate an expectation about the novel action during the test phase without seeing the salient outcome. Thus, they were making use of the learnt association between the action ("touching the toy") and the

---

[8] Furthermore, 6-month-olds did not look longer to the new toy than in the old toy test event when the action-effect was present during both familiarization and test phases, as they should have if they had reacted only to the movement. This age group only interpreted the unfamiliar action in terms of a goal when other action characteristics, such as variations of the approach movement, were also present.

salient outcome ("being picked up") in a new spatial arrangement. Note, however, that the retrieval of this association was taking place in the same general context involving the same objects. It therefore does not prove that infants are indeed able to make use of the learnt link between the novel action and the salient outcome in an entirely new situation. To show that infants are able to fully generalize the particular outcome producing potential of a novel action which they learnt by previously associating the action with a salient outcome, the learning phase has to take place separately from the testing phase and they have to be in different situations.

In the current study, 9- and 12-month-old infants therefore first participated in a training session in which they gained experience with a novel action and its salient outcome. Infants observed the experimenter touching wooden blocks on a table with the back of her hand and picking them up with the help of a Velcro band that she was wearing on her hand. The infants were themselves also given the opportunity to perform this novel action and to produce the interesting outcome. (Note that the aim of the study was not to disentangle the role of observational versus own action experience with the novel action. Our design in the training aimed to allow the infants to gain ample experience about the salient outcome of the action from either of the two sources.) Following the training phase, the infants were moved to another experimental area and participated in a looking time study identical to Woodward’s “back of the hand touch” study (1999). This familiarized them with the back of the hand touching one of two toys without picking it up. After the two toys’ positions had been switched, they were shown two test events in which the back of the hand touched either the same or a new toy.

If the salient effect (pick up) during the training phase enables infants to interpret the novel action (touch with the back of the hand) as being goal-directed, and if they were then able to rely on this interpretation in a new situation even without witnessing the salient effect, then they should be able to generate a goal-directed expectation about this novel action in the test events. Therefore, we expected infants to look longer in the new toy test event than in the old toy test event. Two control conditions were also included to make sure that it is indeed the presence of the action effect that causes the difference in the subsequent interpretation of the novel action in the new situation. One of the control conditions differed from the experimental condition only in that no salient effect was produced during the training phase. That is, the touch of the back of the hand did not result in the picking up of the wooden blocks. In the other control condition, no training phase was included. Infants participated only in the looking time study. We expected that infants would not be able to generate a particular expectation about the target of the action in the two control conditions and that they would thus look equally long in the two test events.

Method

Participants

Twenty-four 9-month-olds (11 males and 13 females, mean age = 39.1 weeks, *SD* = 1.36 weeks, range from 37.1 weeks to 41.6 weeks) participated in the experimental condition. Seventy-two 12-month-olds (40 males and 32 females, mean age = 52.8 weeks, *SD* = 2.3 weeks, range from 47.7 weeks to 56.4 weeks) were assigned into three groups: experimental (24), control 1 (24) and control 2 (24). An additional 9 nine-month-olds and 15 twelve-month-olds were excluded due to fussiness (5, 5), experimenter error during the training or test session (4, 10, respectively). The infants were recruited through mailings or advertisements.

Apparatus

Training session

The infant sat at a table in his/her parent’s lap in the infant lab. The experimenter sat at the other side of the table facing the infant. Four plain wooden blocks (a cube, two cylinders, and a brick varying in size from 8 cm^3 to 32 cm^3) were used, each had a piece of red Velcro attached to its top. The experimenter had a red Velcro band (about 4 cm wide) on her right hand. The infant was wearing a mitten with red Velcro pieces sewed on the back of it[9]. See Figure 1. The pre-training session (which took 3 minutes on average) was video-taped for later scoring. The training sessions for control group 1 and the experimental group were identical except that for control group 1 the Velcro pieces on the wooden blocks did not stick to the band or mitten (because the same "soft" side of Velcro was used on both the blocks and the band/mitten).

---

[9] The idea of using a mitten with Velcro came from the study of Needham et al., 2002 that investigated if sticky mittens can enhance object exploration skills in pre-reaching infants.

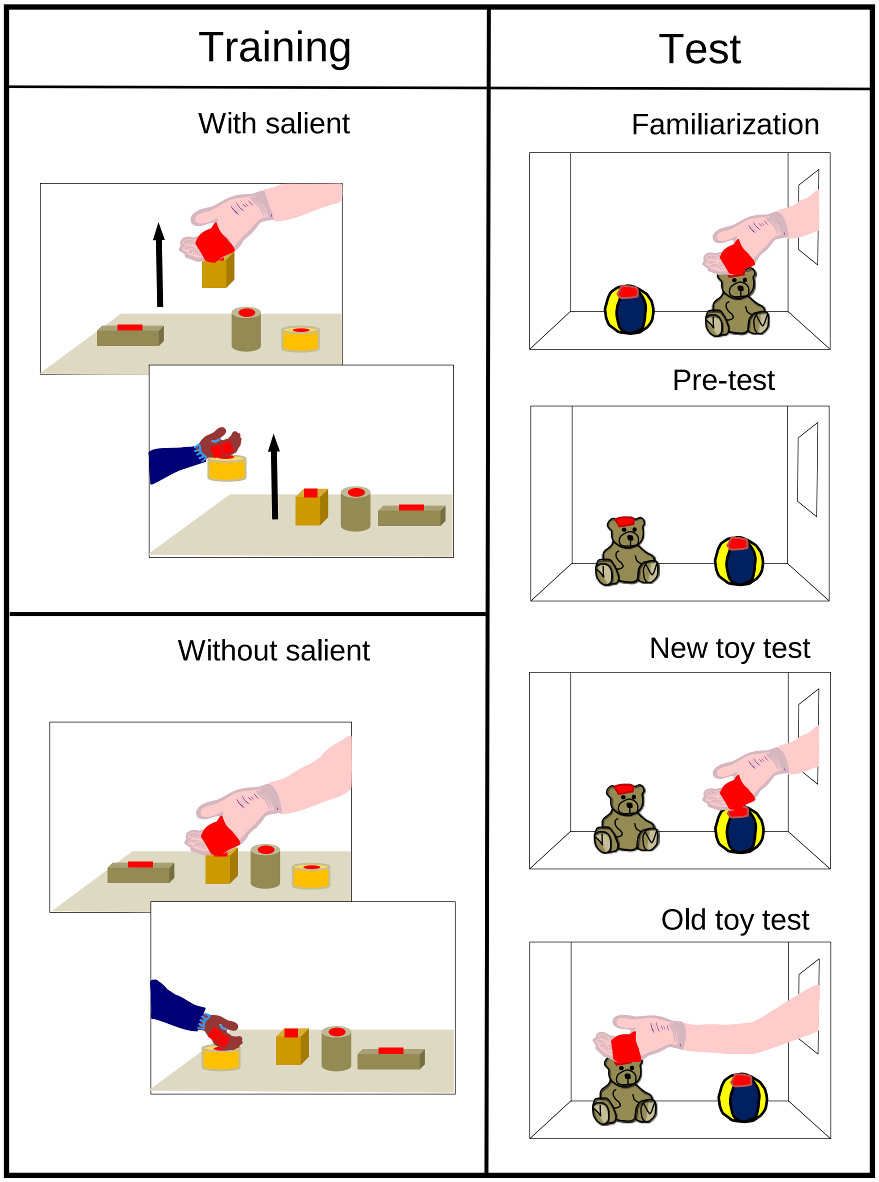

Figure 1. Illustration of training sessions and the subsequent testing session.

Test session

Infants sat on their parent's lap in a darkened and curtained booth that was situated in another part of the infant laboratory room. Infants were facing a stage from a distance of approximately 70 cm. The stage was 60 cm high x 85 cm wide x 40 cm deep and was made from white cardboard. A curtain could be raised to block the stage from view between trials. An opening on the right side of the stage, which was covered by white curtain, allowed the experimenter to enter her arm into the stage area. Two concealed lamps illuminated the stage. A computer signal could turn the lamps on and off to inform the experimenter to start and stop the trials, respectively. A video camera focusing on the baby’s face was mounted above the stage peeping through the opening of a dark curtain. A hidden observer, blind to the test condition and experiment type, timed the infants' looks in the test trials by using a specially built computer program at the displays from head-and-shoulders en face video image. Another camera mounted above and behind the head of the infant also recorded the stage area (a split screen recording was made of all infants, with the stage image on the top and, the baby’s face on the bottom part of the display).

Procedure and Stimuli

Training session for the experimental group

After the parent with the infant sat down at the table, the experimenter placed the first wooden block in front of her on the table and drew the infant’s attention to it. Then the experimenter dropped the back of her hand on the wooden block and, thanks to the Velcro band she was wearing, her hand stuck to the block. Then she lifted up the block. She made sure that the infant was watching and focusing on her action by saying: "Look at my hand!". After this, she separated the wooden block from the band and repeated the same action. As the next step, the experimenter or the parent secured the mitten on the infant’s hand (on the preferred hand if there was such) and the experimenter placed the wooden block in front of the infant. The infant then was encouraged by the experimenter and the parent to imitate the same action with his/her mittened hand (“Do you want to do this too?” or “It's your turn now!”). Parents were instructed to help infants – if necessary - perform the action by gently pressing the back of the mittened hand of the infant against the wooden block. The infant or the experimenter removed the block from the mitten and the infant was offered to pick up the same block once more. The experimenter then hid that block under the table and placed the next block in front of her on the table. The same sequence was repeated with all four blocks. After the training session, the infant and the parent were immediately asked to move to the curtained booth for the test session. The experimenter was the same person for both sessions.

Training session for the control group 1:

The procedure was equivalent to that of the experimental group, with the exception that the experimenter could not pick up the wooden blocks by the band with the back of her hand and thus she simply touched them (since the Velcro pieces did not stick together) and then lifted her hand. Similarly, infants could only touch the wooden blocks with their mittened hand but could not pick them up.

Test session for both the experimental and control groups:

Familiarization phase:

Infants saw four identical familiarization trials. At the start of each trial, the experimenter lowered the curtain to reveal the stage. On the stage floor, there were two toys of around the same size: a yellow plastic Winnie-the-Pooh bear and a ball with green and white patterns. A piece of red Velcro was attached to the top of both toys. After the online observer turned the lights on above the stage using a computer signal, the experimenter reached in through the opening on the right side of the stage holding her palm up. She was wearing the same Velcro band as in the training session. She dropped her hand on top of one of the two toys and stayed still for 15 seconds. (Note the hand did NOT pick up the toys.) See Figure 1. When the end of the 15-second interval was indicated to the experimenter by the light switching off, the experimenter raised the curtain. The position of the toys and location of the action were counterbalanced in the familiarization trials.

Test trials

After the fourth familiarization trial the positions of the two toys were switched behind the curtain. Then the curtain was lowered and the changed positions without the presence of the hand were shown to the infants for 5 sec. Following this short pre-test trial, one of two types of test events was presented to the infants. Half of the infants saw two identical tests trials in which the back of the hand dropped on the same toy as in the familiarization (old toy test event). The other half of the infants saw two tests trials in which the back of the hand dropped on the new toy (new toy test event). In both types of test trials, after the hand dropped on the toy it stayed still as long as the infants watched the scene. (Note that the hand did NOT pick up the toys during the test phase either.) See Figure 1. The online observer started to measure the looking times when the hand touched the toy and if baby was looking at the stage. The test trial ended when infants had looked continuously for a minimum of 2 sec and then looked away continuously for 2 sec. If infants looked away before 2 sec had elapsed, then the trial was ignored and started again. The experimenter raised the curtain when the end of a valid or ignored trial was indicated to her by the light switching off. Parents were asked to close their eyes during the test trials and instructed to refrain from talking to the infant during the whole session except for giving comfort when necessary.

Reliability coding and scoring of the training session:

To assess reliability in the test session and measure behavior in the familiarization phase, an offline observer measured each infant's looking times from the videotaped record. The offline observer was also unaware of the type of the test trials and the experiment type. Inter-observer agreement was computed for each infant's looking times in the two test trials and the agreement was accepted if it was above or equal to 95%. (This was the case for 70% of the cases.) For the rest of the cases a second offline observer was asked to measure the looking times again independently. If this did not lead to agreement with one of the other observers, then the looking times of that infant were excluded from the analyses (this happened in two cases).

In the training session, two offline observers independently scored the infants' behavior from the videotaped records. Infants' observational experience was scored by the absolute duration of their attention to the action demonstration by the experimenter. Since the action demonstration in the experimental groups lasted somewhat longer than in control group 1 due to the fact that the experimenter had to separate the block from the Velcro band, we also calculated the duration of attention relative to the length of the demonstration (duration ratio). Infants' own action experience was scored by the number of times infants properly performed the action. Since turning the hand outward is an awkward and difficult movement for infants, any attempt that resembled such a movement was counted (this was only an issue in the control condition because there was no pick-up). The number of any other actions related to touching the blocks, such as grabbing them with the other hand, banging or shoving them, shaking the mittened hand with the block stuck on, or separating the block with the other hand from the mitten, was also scored. The sum of these two scores was also calculated to obtain an indication of the general level of activity of the infants. Furthermore, the duration of infants' own acting and the latency to start to act were also measured. Parental help was scored by the number of times they helped perform the action and by the duration of parent-assisted action. In addition, we also scored whether infants looked at their hand while their parents moved it and then we calculated the ratio of attended parent-assisted actions by dividing the number of attended parent-assisted actions by the total number of parent-assisted actions. Inter-observer agreement was computed for each score and the agreement was between 0.71 and 0.88.

Test session

Infants' looking times during the familiarization trials were analyzed first. Univariate ANOVA showed no significant difference between the experimental and control groups for the 12-month- olds, $F(2,68) = 2.44$, $p = .10$, $\eta_p^2 = .07$, or between the 9 and 12 months old experimental groups, $F(1,46) = .57$, $p = .45$, $\eta_p^2 = .01$, in the average looking times of the four familiarization trials. Infants in all groups showed a decline in their looking times across the familiarization trials, the difference between the 1st and 4th familiarization trials were significant in all groups, 9 months: $t(23) = 5.45$, $p < .001$; 12 months experimental: $t(23) = 5.17$, $p < .001$; 12 months control 1: $t(23) = 4.25$, $p < .001$; 12 months control 2: $t(23) = 4.24$, $p < .001$. The looking times in the pre-test trial in which the infants could see the changed position of the toys on the stage did not differ between the experimental and control groups for the 12-month-olds, $F(2,68) = 2.51$, $p = .09$, $\eta_p^2 = .07$, or between the 9 and 12 months old experimental groups, $F(1,46) = .89$, $p = .35$, $\eta_p^2 = .02$, see Figure 2.

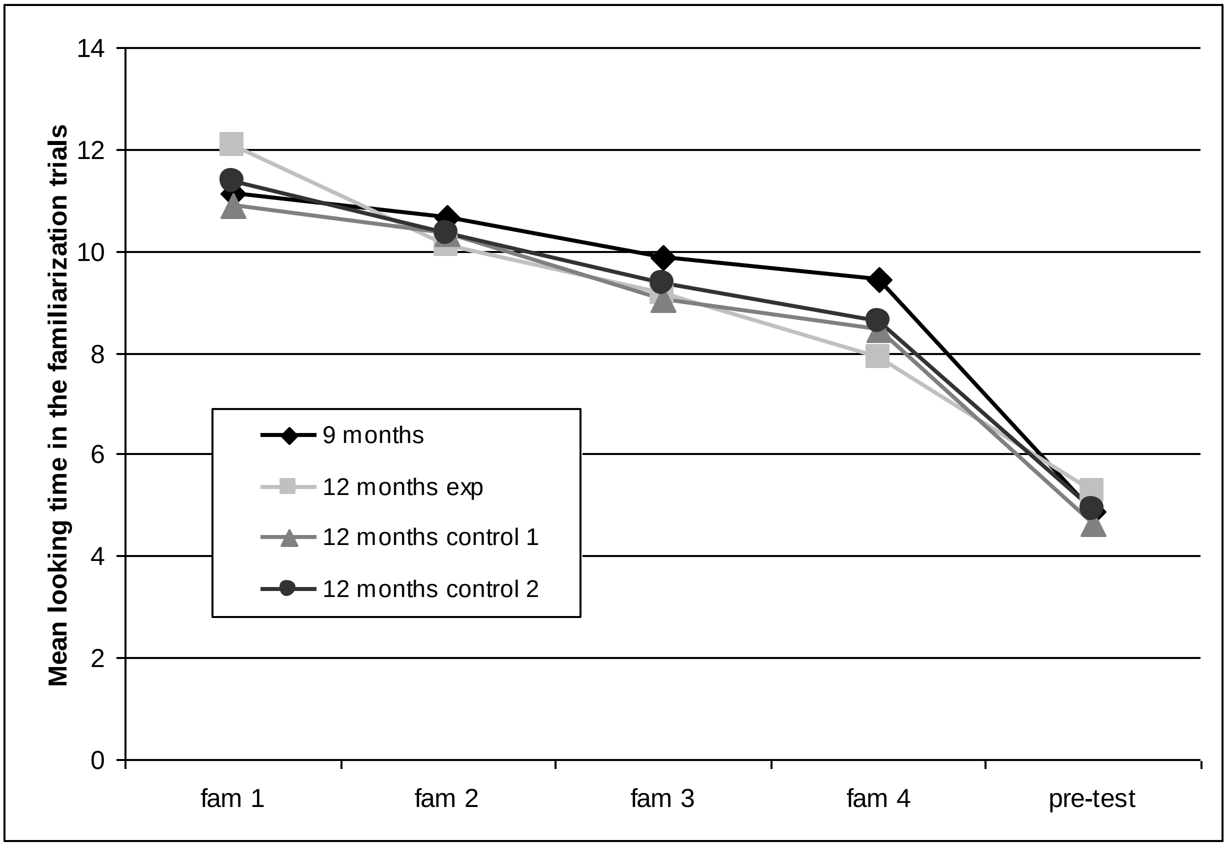

Figure 2. Mean looking times during the familiarization trials for the 9- and the three 12-month-old groups. (The lengths of the first four trials were 13 sec, the pre-test trial was 5 sec long.)

The looking times of the two test trials were analyzed separately for the two age groups. Preliminary analysis of looking times for the two test trials did not reveal any effect of gender, location of the action (near side vs. far side), type of toy touched (bear vs. ball) or the position of the toys (bear left vs. bear right) in any of the groups. Therefore, these factors were omitted from further analyses. Using test type (2, new toy and old toy) and experiment type (3, experimental, control 1 and 2) as between-subject factors and trial (2, first and second) as a within-subject factor, a repeated measures ANOVA was carried out for the 12 months old infants. The analysis found a main effect of trial, $F(1,66) = 4.99$, $p = .029$, $\eta_p^2 = .07$, indicating that infants overall looked longer in the 1st than in the 2nd test event. A significant effect of trial x test type x experiment type interaction, $F(2,66) = 3.74$, $p = .029$, $\eta_p^2 = .10$, was also revealed[10]. To explore this interaction, separate ANOVAs were carried out in each experiment type group. The ANOVAs showed no main or interaction effects of trial or test type factor in control group 1, control group or 2. On the other hand, in the experimental group a trial and test type interaction was found, $F(1,22) = 6.13$, $p = .021$, $\eta_p^2 = .21$. T-tests showed that infants looked longer in the new toy test event than in the old toy test event in the first test trial, $t(22) = 2.54$, $p = .019$, two-tailed, but not in the second test trial, $t(22) = -.74$, $p = .46$, see Figure 3. These results were confirmed with non-parametric tests (first test trial in the experimental group: Mann-Whitney $U = 33.5$, z = -2.22, p = .024, two-tailed). In addition, further analyses revealed that during the 1st test event infants who saw the old toy test event in the experimental group looked significantly less long, $F(2,34) = 4.80$, $p = .014$, $\eta_p^2 = .22$, than infants who saw the same test event in the control 1 ($p = .046$) or control 2 group ($p = .005$). There was no difference between the three groups in the looking times of infants who saw the new toy test event or during the 2nd test event in any of the test types.

---

[10] These effects remained the same when the difference between the looking times for the 1st and 4th familiarization trials was included as a covariate in the analysis. Thus, the amount of decline had no effect on the looking time pattern in the test events.

In the 9 months old group a repeated measures ANOVA for looking times in the test events with trial (2) as within-subject factor and test type (2) as between-subject factor revealed only a trial effect, $F(1,22) = 7.23$, $p = .013$, $\eta_p^2 = .24$, indicating longer looking times in the first than in the second trial. No effect of test type was found, see Figure 3. We also compared the two age groups in their looking times in the test events by using a repeated measures ANOVA with trial as within-subject factor and test type and age group as between-subject factors. We found a three-way interaction between trial, age group and test type, $F(1,44) = 4.09$, $p = .049$, $\eta_p^2 = .08$. This interaction indicates that - in addition to the looking pattern that has already been reported separately for the two age groups - there was a significant difference between 1st and 2nd test trials in the 9 months old group that saw the old goal test event, $t(11) = 2.85$, $p = .016$, while there was no such difference in the 12 months old group, $t(11) = -1.30$, $p > .21$.

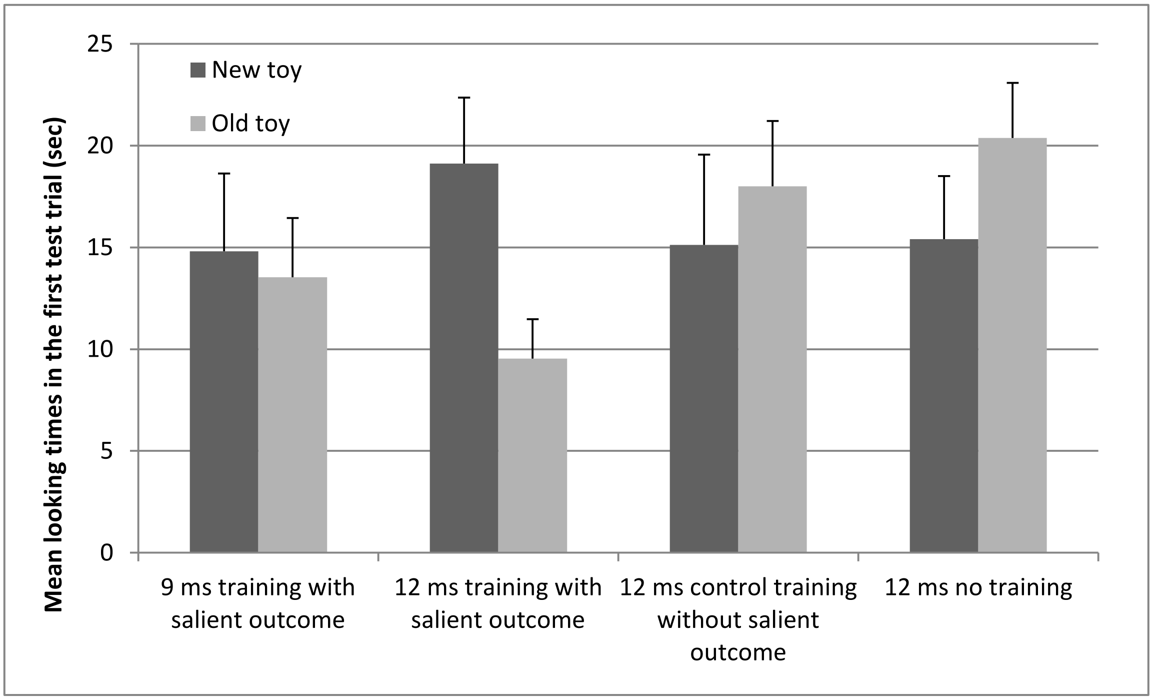

Figure 3. Mean looking times with standard errors in the 1st test trial (sec) by test event type for the 9- and the three 12-month-old groups. * = p<0.05

Training session

Since the demonstration during the training session was somewhat longer in the experimental group than in the control due to the fact that the experimenter had to separate the block from the Velcro band in the experimental group, the absolute time the 12 months old infants attended the demonstration was longer in the experimental group than in the control 1 group condition, $t(46) = 3.72$, $p = .001$] . However, there was no difference between the two groups in the time infants spent attending the demonstration relative to the total time of the demonstration (duration ratio), $t(46) = - 2.25$, $p = .06$ The two age groups in the experimental condition (9 and 12 months) did not differ from each other in absolute attention duration, t(46) = -.66 , $p = .50$, or relative attention duration, t(46) = .91 ,$p = .36$. We found no further differences between any of the groups in the scores for infants' own action experience or parental assistance ($p_s > .20$) except for one of the scores. The *t*-test revealed a difference between the two age groups in the ratio of attended parent-assisted action, $t(46) = -2.89$, $p = .006$, indicating that the 9 months olds were paying significantly less attention to their own hand while their parents moved it than the experimental 12 months old group did. See Table 1 for a summary of the training scores.

Next, we checked whether there was any difference in age, looking times during the familiarization trials or in the training scores between infants who were allocated to the group that saw the New toy or the Old toy test event. No difference was found in any of these measures in any of the groups ($p_s > .15$). Finally, to examine whether the behavior of the infants during the training session had any effect on their looking patterns in the subsequent test session, all the analyses of variance of looking times for the test events were also carried out with all training scores entered as covariates (ANCOVAs). The pattern of results remained unchanged in all analyses.

Table 1. Summary of training scores.

	Attention to demonstration				**Infants own action**										**Parental assistance**					
	Absolute duration of attention		Attention duration ratio		Number of Proper action		Number of Other action		Sum of action		Duration of action		Mean Latency to act		Parent action duration		Number of Parent-assisted action		Attended parent-assisted action ratio	
	Mean	s.e.	Mean	s.e.	Mean	s.e.	Mean	s.e.	Mean	s.e.	Mean	s.e.	Mean	s.e.	Mean	s.e.	Mean	s.e.	Mean	s.e.
9 month experimental	47.73	3.42	.97	.01	1.71	.43	4.67	.94	6.38	1.10	30.24	5.52	2,14	0.51	19.10	2.57	4.92	.44	.74	.05
12 months experimental	51.11	3.71	.96	.02	1.33	.32	4.79	.77	6.13	.86	23.59	3.47	2.83	0.66	15.58	1.69	4.83	.56	.91	.03
12 months control 1	36.24	1.47	.98	.01	1.04	.28	5.08	.57	6.13	.60	37.02	6.08	3.14	2.36	25.50	9.37	4.79	.45	.81	.07

Discussion

We tested the ability of 9 and 12 months old infants to interpret a novel action as goal-directed when they previously could associate this novel action with a salient effect in another situation. We found that 12 months old infants who took part in the experimental training session in which they were provided experience with a novel action accompanied by a salient effect, looked longer in the new toy test event than in the old toy test event in the subsequent testing session. This result thus suggests that the infants were able to generate a particular expectation about the goal of this novel act in a new situation even in the absence of the salient action effect. Note that the difference in looking time was only present in the first test trial and not in the second, which is however not surprising or unprecedented as during repeated presentations infants' expectation may change by reinterpreting the situation on the basis of previous test trials. The 12 months olds in the two control conditions in which the novel action did not produce a salient effect, did not show evidence of being able to interpret the novel action as goal-directed in the testing session, as they looked equally long in the two test events. In addition, infants in both control groups looked longer at the old toy test event in the first test trial than infants in the experimental group which further supports the interpretation that only the experimental group expected this event to take place. In the 9 months old group there was no difference in looking times between the two test events, which indicates that the training session with salient effect did not facilitate a goal-directed interpretation of the novel action for these infants. Taken together we demonstrated that, by 12 months of age, infants can generalize the outcome producing potential of a novel action to interpret the novel action as goal-directed.

The performance of the 12 months olds in the two control conditions clearly indicates that it was the presence of the salient action effect in the experimental condition that was critical for changing the subsequent interpretation of the novel action. The lack of influence of the training session in control condition 1 excludes the possibility that the facilitating effect of the training in the experimental condition was achieved simply by receiving more prior (observational) experience with the unfamiliar action or by seeing this action being performed by a person who was interacting with the infants and who drew their attention to the action. The lack of difference between the looking times for the test events in control condition 2 shows that the original "back of the hand touch" action without effect is not interpreted as goal-directed not only by 6- and 9- as had previously been demonstrated (Woodward, 1999), but also by 12-month-olds. Furthermore, this latter result also excludes the possibility that the difference between the experimental and control 1 conditions reflects disruptive effects of the control training rather than a facilitating effect of the experimental training.

Our study suggests that a) a salient outcome is critical in interpreting a novel action as goal-directed and that b) 12 months old infants can make use of a learnt link between the action and its outcome in a new context by interpreting the novel action in terms of a goal even in the absence of the salient outcome. These findings can be viewed as a demonstration of cue-based bootstrapping by which infants' initial sensitivity to behavioral cues is united with a learning process about different types of actions (Biro & Leslie, 2007).

Why did the 9 months olds not benefit from the training? One simple answer could be that younger infants would have needed more experience with an unfamiliar action and its effect in order to link them together. Another possibility is that the infants could link the novel action to the outcome, but had difficulties with retrieving the learnt association. Studies on imitation and memory of own action-effect contingencies show that contextual changes such as changes in the room surroundings or cues can greatly influence the quality of imitation or the memory of actions (Rovee-Collier, 1996; Moore & Meltzoff, 2004). Since our training and test sessions took place in different surroundings, contextual generalization was certainly necessary. It has been suggested that the ability to generalize across contexts develops somewhere around 9 months of age. However, the exact age at which infants can transfer certain information seems to depend on the degree of change and on the task involved, which makes different studies difficult to compare. In the domain of action understanding, Sommerville and her colleagues (2009) recent study is the most relevant one for our purposes. They found that 10 months olds could only use prior information about the particular goal of an action to generate an expectation about a means action if the information was given in the same room. Imitation studies also show that, by 12 months, infants are not affected by context change in their recall of observed actions (e.g., Klein & Meltzoff, 1999).

Besides contextual differences, the difficulty with the retrieval and generalization of the associated link may also have been caused by the manner in which the infants encoded the observed action during the training session. As we have argued, an important step in learning a new goal-directed action is to be able to determine which aspects of the action are causally necessary to achieve the outcome. For example, if the infants assumed that the person as a whole, her other movements, her smile, or the particular objects that were used were (also) relevant to produce the salient effect in the training session, then the absence of these features in the test session could have prevented the infants from making use of the association. Thus, the retrieval/generalization problems may not (only) stem from the way memory for events are organized, but also from the difficulties with applying teleological reasoning due to the limited prior experience and knowledge about actions and actors at this age.

It is also important to differentiate our study from other training studies with similar features. Elsner and Ashcerleben (2003), for example, investigated the influence of salient action effects produced by a novel object by providing observational training to 9-18 months old infants. They found that only from the age of 12 months did infants benefit from a salient action effect. What they tested, however, was not whether the effect influenced infants' goal-directed interpretation of an observed novel action but whether it changed their own exploratory behavior. Nevertheless, their developmental pattern fits with our finding.

Hofer and her colleagues (2005), on the other hand, tested 9 months old infants' goal-directed interpretation of a novel tool by providing a short prior demonstration in which the tool was handled by both a model and the infants themselves. Here, however, the subsequent test session did provide salient action-effects. Their training study therefore did not aim to test whether infants can interpret a novel action as goal-directed in the absence of the outcome.

In two training studies, one that provided observational experience with the particular goal of an actor (Sommerville & Crane, 2009) and one that provided own action experience with a novel means-end action (Sommerville et al., 2008), it was found that 10 months old infants had benefited from these trainings when they were subsequently tested on their understanding of observed means-end relationships. However, the entire means-end sequence was presented in the habituation phase of the testing, thus the infants were not required to solely rely on the association between the action and the outcome that they had built up during the training.

Sommerville, Woodward and Needham's (2005) study is the most similar to the current training study. In their study 3.5 months old infants, who cannot reach and grasp objects themselves, participated in a training in which they were wearing a mitten with Velcro attached to the palm and with which they could pick up two toys. This experience allowed them to interpret a grasping hand of another person (without the salient effect of pick-up) as goal-directed. How can we reconcile the large age difference between that study and ours? One possibility might be that in Sommerville, Woodward and Needham's study the objects that the infants picked up during the training were identical to those that were used in the subsequent habituation paradigm. Hence, there was less demand on the ability to generalize the outcome producing potential of the mittened grasping hand from training to test. Alternatively, while grasping and then transferring an object is a natural hand action with which even young infants have ample prior observational experience, picking up objects by touching them with the back of the hand is unnatural, violates infants' existing physical knowledge (about support and mechanics) and may thus be harder to learn.

A further possibility to explain the difference between these studies leads us to the question of the role of the type of experience that infants can benefit from to interpret an observed novel action as goal-directed. Note, however, that the current study was not designed to test whether the interpretation of an observed action is primarily influenced by infants' own experience with the action or by observational experience. Instead, we emphasized the importance of appreciating the power of an action to produce outcomes in infant's goal-directed action interpretation. We therefore varied the presence or absence of a salient effect via both types of experience.

A recent study that has contrasted the impact of infants' own action experience and observational experience on the interpretation of means-end actions, found self-experience to be primary (Sommerville et al., 2008). However, in both this study and in those that show a strong correlation between infants' own ability to produce certain means-end actions and their understanding of others' similar actions in terms of goals (Sommerville and Woodward, 2005a, 2005b), a salient action outcome was always produced by the infants' own action. Our study is the first to show that, without this salient outcome, own action experience with the novel action does not enhance the ability to interpret the novel action as goal-directed. The salient outcome thus seems to be the crucial mediating factor.

Another reason for the age difference between our study and Sommerville, Woodward and Needham's (2005) might therefore be that the amount of own action experience with the novel action and its effect was insufficient for the 9 months olds in our study. The coding of our training session shows that infants indeed scored low on the number of proper own action scale. This might have been due to difficulties with imitating the awkward back of the hand touch action or due to a dislike of wearing the mitten. On the other hand, the 12-month-olds scored just as low on this scale, which suggests that either even a small amount of self experience was sufficient for them or that they could make better use of the observational experience by this age. We found, however, that, compared to 12-month-olds, 9-month-olds looked less at their own moving hand while their parents helped them to produce the novel action with its effect. It is therefore possible that 9-month-olds indeed gained somewhat less experience with the effect of the novel action from the training session due to their relative lack of attention during parental help. Future research, however, is required to clarify the exact role of observational vs. motor experience in learning about novel goal-directed actions.

Finally, it is important to point out that in the current study it is hard to distinguish whether infants learnt a new means action or the function of a novel tool, or both. In other words, during the training infants may not (only) have associated the salient effect with the “back of the hand touch” action, but (also) with the Velcro band itself. That is, the infants may have learnt that the Velcro band is *for* picking things up. In the testing phase, the hand was presented with the Velcro band on and the infants could thus rely on the inferred function of the Velcro band in their interpretation of the action[11]. Future research can for example disentangle these possibilities by using a Velcro band with two different types of hand actions during the test and training phases.

The function of a tool is usually learnt together with a certain action by which the tool can be used to produce an outcome. While for certain tools the proper (i.e. most efficient) use requires very

---

[11] This could also have added to the possible generalization difficulty of the 9 months olds, because the mitten they had self experience with and the Velcro band they observed did not look exactly the same.

specific actions (e.g. a pair of scissors), for other tools the range of actions that allows proper functioning is much wider (e.g. a towel). An interesting question is therefore what types of cues help infants to figure out the level of specificity of the action by which a tool can best be operated.

In summary, we have shown that experiencing a salient effect of a novel hand action (or use of a novel tool) can alter 12 months old infants' interpretation of the action (or the tool). By appreciating the potential of the action (or tool) to produce an outcome, infants can attribute a goal when they observe the novel action (or novel tool) in a new situation. Our finding shows the essential role a salient outcome plays in the developing ability of infants' goal-directed understanding of novel actions.

Chapter 5: Explicit learning of arbitrary and non-arbitrary action-effect relations in adults and 4-year-olds

Stephan A. Verschoor, Rena M. Eenshuistra, Jutta Kray, Bernhard Hommel

Chapter 5

Abstract

Ideomotor theories claim that carrying out a movement that produces a perceivable effect creates a bidirectional association between the two, which can then be used by action control processes to retrieve the associated action by anticipating its outcome. Previous implicit-learning studies have shown that practice renders novel but action-contingent stimuli effective retrieval cues of the action they used to follow, suggesting that experiencing sequences of actions and effects creates bidirectional action-effect associations. We investigated whether action-effect associations are also acquired under explicit-learning conditions and whether familiar action-effect relations (such as between a trumpet and a trumpet sound) are learned the same way as novel, arbitrary relations are. We also investigated whether these factors affect adults and 4-year-old children equally. Findings suggest that explicit learning produces the same bidirectional action-effect associations as implicit learning does, that non-arbitrary relations improve performance without affecting learning per se, and that adults and young children show equivalent performance—apart from the common observation that children have greater difficulty to withstand stimulus-induced action tendencies.

Chapter 5

Introduction

James' (1890) ideomotor theory claims that consciously thinking of an action goal automatically triggers the accompanying actions that will help to reach that goal. This ideomotor approach has proved to be very useful in interpreting and explaining goal-directed behavior. For example, if you intend to watch your favorite sitcom, it would be of good help if actions such as finding the remote control, switch on the TV and looking for the right channel already get activated in order to accomplish your goal smoothly and correctly. If no actions get primed if an action plan comes to mind or if incorrect actions get triggered, such as getting a spoon or opening a door, reaching your goal would be severely disturbed or delayed. As the example shows, planning and control of goal-directed actions does not only require knowledge or expectations about the outcomes of these actions, it also implies that the relationship between the action and the action outcomes has to be *bidirectional* in order to use action-outcomes as a trigger for action initiation: even though acquiring an action-effect relation implies that the cognitive representation of the given action gets activated before the representation of this action’s effects (as the action precedes its effects), planning that action later on requires the representation of the effect to get activated before the representation of the action.

Ideomotor approaches to action planning and action control recently regained interest (Elsner & Hommel, 2001; Hommel, 1996; Hommel et al., 2001; Stock & Stock, 2004) and a number of studies have demonstrated that subjects indeed acquire and use bidirectional action-outcome contingencies to plan and guide their actions (for an overview, see Hommel & Elsner, 2009). Most of these studies adopted the experimental design introduced by Hommel (1996; Elsner & Hommel, 2001), which comprises of two phases: an acquisition phase, in which an action-outcome contingency is learned, and a subsequent transfer phase, in which it is tested whether the acquired action-outcome contingency affects action control.

In an acquisition phase of this design subjects freely choose among several actions (say, the key-pressing response R1 or R2) and each action is consistently followed by an effect stimulus (say, sound E1 or E2). The effect stimulus is not relevant to the task and subjects are not encouraged in any way to attend to, or acquire the action-effect relationship. In the transfer phase the effect stimuli are now being presented as target stimuli, and the sound-key mapping can be either consistent (E1 → R1 and E2 → R2) or inconsistent (E1 → R2 and E2 → R1) with the key-sound mapping of the acquisition phase. In general, it is found that subjects perform better during the transfer phase when an acquisition-consistent effect-action mapping is required as compared to an inconsistent mapping (e.g., Elsner & Hommel, 2001). Meaning that subjects respond faster and/or make fewer errors during the transfer phase when asked to respond to effects with the keys that previously caused these effects during acquisition, than when this mapping is reversed from that acquired during acquisition. This effect of mapping consistency on RTs and errors suggests that, despite the task-irrelevance of the action effects and the action-effect relationship during acquisition, bidirectional action-outcome associations (R1 ↔ E1 and R2 ↔ E2) were created during acquisition and this knowledge transfers to the transfer phase. Subjects with an acquisition-consistent mapping during the transfer phase apparently benefited from the action-outcome contingencies learned during the acquisition phase and use this knowledge to guide their actions in the transfer phase. It is also possible that performance on acquisition-inconsistent mappings was hampered by having learned inconsistent contingencies, but there is some evidence that positive transfer effects by far outweigh negative effects (Hommel, 2004).

The available studies on action-effect learning have used novel, arbitrary relations between action effects, so that the amount of learning could be experimentally controlled. Moreover, as the novel action effects are commonly task-irrelevant and subjects are often explicitly instructed to ignore them, the acquisition of action-effect associations must be considered implicit or at least non-intentional. Indeed, Elsner and Hommel (2004) were unable to find any relationship between the reported experience of a causal relationship between action and effect in the acquisition phase and the size of the consistency effect in the test phase. Verschoor et al. (2013) recently showed that no awareness of action-effect relationships is required at all to nonetheless show action-effect learning in a transfer phase.

Even though it is theoretically important to demonstrate that spontaneous action-effect learning can take place under such “unfavorable” circumstances, one may argue that these circumstances are not particularly ecologically valid. Infants, children, and adult novices facing a new task will often actively explore the appropriateness and potential of alternative actions to reach a particular goal, and thus explicitly carry out specific actions to produce specific effects. In the present study, we investigated whether these circumstances also allow for the acquisition of bidirectional action-effect associations. Apart from ecological validity considerations, this research goal was motivated by some recent observations that seem to call into doubt that intentional action-effect learning leads to bidirectional associations.

As Verschoor, Eenshuistra, Weidema and Hommel (submitted) demonstrated, requiring subjects to verbalize the causal relation between actions and effects eliminates the consistency effects in the transfer phase of the Elsner and Hommel (2001) paradigm. This might suggest that consciously representing causal relations emphasizes their unidirectional nature and thus either prevents the creation of bidirectional associations between action representations and effect representations or blocks them from impacting action control (Verschoor et al., submitted). A similar line of reasoning is suggested by more general views about associative and causal relations between events. While associative relations are supposed to be bidirectional and thus insensitive to the temporal sequence between the represented events, causal relations are usually seen as asymmetrical and irreversible (Hausman, 1998; Hume, 1739/1964; see Waldmann, 1996). Consistent with Verschoor et al. (submitted), Arcediano, Escobar and Miller (2005) suggested that the use of verbal labels may stress the temporal relation between two events and as consequence could constrain the direction of the association. A study of Fenker, Waldmann and Holyoak (2005) also attributed the difference between associative and causal relations. When subjects had to judge if sequences of two words (e.g., spark → fire or fire → spark) were presented in a causal (cause → effect) order or in a diagnostic order (effect → cause), they were much faster in judging the causal relations than the diagnostic relations. However when subjects were asked to judge if the same word sequences had an associative relation (spark ↔ fire or fire ↔ spark), the temporal order was of no importance anymore. It thus makes sense to assume that relations between actions and effects can be represented in both a bidirectional, associative manner that previous studies on implicit action-effect learning seem to have tapped, and a unidirectional, causal way that emphasizes, or is a consequence of emphasizing, the temporal sequence of actions and effects.

The question of this study thus was whether explicit action-effect acquisition would be sufficient to eliminate bidirectional associations, which would put rather tight constraints on ideomotor theorizing in the context of action control, or whether explicit acquisition would be comparable to implicit acquisition as measured by many other studies on action-effect learning (e.g., Elsner & Hommel, 2001). In the acquisition phase of the present experiment we made up a cover story that explicitly related key-pressing actions to auditory effects in a causal fashion. The two response keys were introduced to symbolize (and labeled by pictures of) two sound-producing objects, a trumpet and a bell. Pressing the key was described as acting upon the respective object and subjects were instructed to make the musical instruments sound by pressing the corresponding keys. Consequently, subjects were made aware of the action, the outcome, and the sequential and causal relation between the two. If indeed, as Verschoor et al. (submitted) claim, explicit acquisition of action-effect knowledge is sufficient to render action-effect associations unidirectional, we would expect no effects of mapping on either RTs or errors.

In an additional attempt to consider more realistic situations, we were also interested to see whether pre-existing knowledge about the causal relationship between an action and its effects would alter the directionality of the respective association. To investigate that, we compared learning of a novel, unfamiliar, and arbitrary pairing of a key-pressing action with a tone, with learning of a pairing that we considered familiar to our subjects from pre-experimental experience (i.e., non-arbitrary, as defined in reference to that experience). To accomplish this contrast between non-arbitrary and arbitrary action effects, we manipulated the actions effects so that one key would produced the sound that the respective object was known to produce (e.g., pressing the trumpet key produced a typical trumpet sound, from here on symbolized as <toot>), whereas pressing the other key produced a novel and arbitrarily chosen sound (e.g., bell key → <piew>). There are at least two possible outcomes of this familiarity manipulation (a further possibility will be discussed later). First, one may assume that pre-existing associations facilitate bidirectional learning, so that the familiar relationship produces a stronger consistency effect. Second, however, it may be that knowing about the causal relationship between trumpets and the sound they produce strengthens the association from trumpet to <toot> but works against creating a bidirectional association. This should reduce or eliminate the consistency effects on RTs and error rates in the transfer phase for the non-arbitrary action effects.

A final purpose of the present study was to compare the learning performance of adult subjects with 4-year-old children. In previous studies we obtained evidence of action-effect acquisition in both adults and children, even though the performance profiles differed in detail (Eenshuistra et al., 2004; Kray et al., 2006): Whereas RTs showed equivalent mapping-consistency effects in adults and children (with acquisition-consistent mappings yielding faster RTs), error rates exhibited considerably stronger consistency effects in children (with greater error rates in acquisition-inconsistent mappings). Indeed, in young children facing a just-acquired action effect seems to prime the associated action so strongly that they find it extremely difficult to impossible to select a different action if the instructions so demands—an observation that Eenshuistra et al. (2004) considered an example of goal neglect in the sense of Duncan (1995). Considering this systematic difference between adults and young children, we were interested to see whether comparable differences would be obtained under conditions that make the causal relation between actions and effects explicit, as in the present study, and whether possible effects of pre-experimental knowledge would affect adults and children in the same fashion.

Method

Participants

Participants were 34 4-year-old children (*M* = 4.54, *SD* = 0.37) recruited from two kindergartens in the Netherlands and 35 undergraduate students from the Leiden University (*M* = 21.86, *SD* = 3.29). The data of two children and three students were excluded from the analysis because they were unable to perform the tasks and did not adhere to the task instructions, respectively. The children received small presents for participating, and the kindergartens received book tokens. Students received course credits or €3.50 for participating.

Tasks and Stimuli

The experiment was divided into an acquisition phase and a transfer phase. Trials in the acquisition phase started with the presentation of a visual stimulus, a picture of Ernie or Bert from Sesame Street. A picture of Ernie signaled a go trial, which required a freely chosen left or right key press on a specially designed two-key keyboard. One key was marked with a picture of a trumpet and the other key with a picture of a bell. Pressing one of the two keys was always followed by a sound that was plausibly related to the picture marking the key (trumpet key followed by <toot> or bell key followed by <tring>) resulting in a *non-arbitrary* action-tone relation. Pressing the other key led to the presentation of a sound that bore no pre-experimental relation to the picture label of the key (trumpet key and <piew> or bell key and <piew>) resulting in an *arbitrary* action-tone relation. For example, pressing the left trumpet key could be followed by a <toot> sound and pressing the right bell key by a <piew> sound. All key-sound mappings were balanced across subjects and all sounds were presented for 200 ms. Ernie’s picture remained on the screen until a response was given or until 7 s had passed. The next trial started 1500 ms after a response. When a picture of Bert was presented, no response was required (a no-go trial). The picture of Bert remained on the screen for 2 s.

The transfer phase consisted of a go/no-go two-choice reaction-time task. Again a picture of Ernie signaled a go trial and a picture of Bert a no-go trial (no-go trials were used to make the task more appealing). The picture of Ernie staid on screen until a response was made or until 2 s had passed. In a no-go trial a picture of Bert without any sound was presented for 2 s. The inter-trial interval was 1500 ms.

A go trial started with the presentation of the picture of Ernie together with one of the two sounds that in the preceding acquisition phase served as action effects (200 ms). The action effects now served as imperative stimuli. Participants were to respond to the stimuli as quickly and accurately as possible according to a fixed stimulus-response (S-R) mapping. They were randomly assigned to either a consistent- or an inconsistent S-R mapping in the transfer phase. In the consistent mapping condition, the sound-key mapping matched the key-sound mapping of the acquisition phase (e.g., trumpet key → <toot> and bell → <piew> in the acquisition phase became <toot> → trumpet key and <piew> → bell in the transfer phase). In the inconsistent mapping condition, the sound-key assignment was reversed (e.g., trumpet key → <toot> and bell → <piew> in the acquisition phase became <piew> → trumpet key and <toot> → bell key in the transfer phase). As in our previous study (Eenshuistra et al., 2004), each key press in the transfer phase triggered the same sound as in the acquisition phase, to avoid extinction of the action-effect associations. Previous research shows that this does not alter effect patterns on RT & errors in adults (Elsner & Hommel 2001; Eenshuistra et al., 2004) but it greatly increases effect sizes in children. This finding was taken to suggest that action-effect knowledge in children quickly fades during transfer if the effect tones in response to a keypress are absent.

Associatedness was coded with respect to the relation between the key label (trumpet or bell label) and the sound; e.g., responding with the trumpet key to <toot> would fall into the non-arbitrary category while responding with the bell key to <piew> would count as arbitrary.

Procedure

The acquisition phase consisted of 18 practice trials (12 go trials, 6 no-go trials) and 144 test trials (96 go trials and 48 no-go trials). Subjects were instructed that Ernie "likes music" and that they could "make music" for Ernie. They were told that they could choose freely which sound to make for Ernie by pressing the trumpet key or the bell key. The causal relation between pushing one of the keys and the resulting auditory effect was made explicit by instructing participants that "If you want the trumpet to make a sound you should push the trumpet-key" and giving them the opportunity to explore the corresponding contingency by pressing the key. The same instruction followed for the bell key. Additionally, subjects were instructed that they could freely choose how often they pressed a key but that in total they had to press both keys about equally often. When participants persevered in pushing only one button during acquisition (predominantly children were prone to this), the experimenter reminded them to change keys now and then. Furthermore, they were told that Bert "does not like music" and that no response should be given when a picture of Bert appeared. To motivate subjects—especially the children—to complete the acquisition phase, they were told that Ernie was allowed to play until lunchtime. A picture of a clock that was colored for one third and two third was presented after 48 and 96 trials, respectively. After 144 trials the clock was completely colored, indicating that lunch time had reached and with that the end of the acquisition phase.

In the transfer phase participants were told that Ernie was making music and wanted them to participate. When a picture of Ernie appeared together with one of the two effect sounds, they should press the trumpet key and when they perceived Ernie together with the other effect sound they should press the bell key, thus motivating the (consistent or inconsistent) S-R mapping required for the transfer phase. Again, subjects were instructed to withhold their response when Bert appeared, because ‘‘Bert hates music and likes the silence’’. As in the acquisition phase, subjects were informed about their progress in the transfer phase. However, now they were instructed that Ernie was allowed to make music until bedtime and the same clock procedure was used to indicate when bedtime had reached and the transfer phase had finished. The transfer phase consisted of 18 practice trials (12 go trials, 6 no-go trials) and 90 test trials (72 go trials and 18 no-go trials).

Results

Acquisition Phase

Left and right hand responses were equally distributed (49.8% vs. 50.2%, $t(63) = .42$, $p > .65$) and equally fast ($F(1, 60) = .37$, $p > .5$). As expected, adults were substantially faster than 4-year-olds (428 vs. 1215ms respectively), $F(1, 60) = 290.09$, $p < .001$, η2p = .83) . Overall, responses were faster when they were followed by a sound effect that was pre-experimentally associated with the key label than by a sound that was not (805 vs. 839ms respectively), $F(1, 60) = 6.82$, $p < .02$, η2p = .10. This effect of associatedness interacted with age group ($F(1, 60) = 4.75$, $p < .04$, η2p = .07) due to that the associatedness effect was reliable in the children (1185 vs. 1246ms for non-arbitrary and arbitrary relations) ($F(1, 30) = 5.92$, $p < .05$, η2p = .17) but not in the adults (p > .20). No *a priori* effects of consistency were found.

Transfer Phase

RTs. Mean RTs of the transfer phase were analyzed as a function of age group, associatedness (arbitrary versus non-arbitrary), and mapping (consistency versus inconsistent), with mean overall RT of the acquisition phase as covariate. Unsurprisingly, the adults were faster than the 4-year-old children (630 vs. 1120ms for adults and children), $F(1, 59) = 47.41$, $p < .001$, $\eta 2p = .45$. As predicted by the hypothesis of bidirectional action-effect associations, the consistent mapping yielded faster responses than the inconsistent mapping (consistent mapping 801ms inconsistent mapping 950ms), $F(1, 59) = 25.03$, $p < .001$, $\eta 2p = .30$ (see Figure 1). The interaction between age group and mapping consistency was not significant, $p > .5$, suggesting that action-effect learning was comparable in children and adults. Overall, responses were faster to non-arbitrary action-tone relations than to arbitrary relations (873 vs. 879ms respectively), $F(1, 59) = 4.33$, $p < .05$, $\eta 2p = .07$. Like in the acquisition phase, this effect was numerically more pronounced in children (1096 vs. 1145ms for non-arbitrary and arbitrary relations, respectively) than in adults, where it was reversed in sign (650 vs. 612ms), but the interaction between associatedness and age was not reliable, $p = .19$. All other interactions with associatedness also failed to reach significance ($p > .5$).

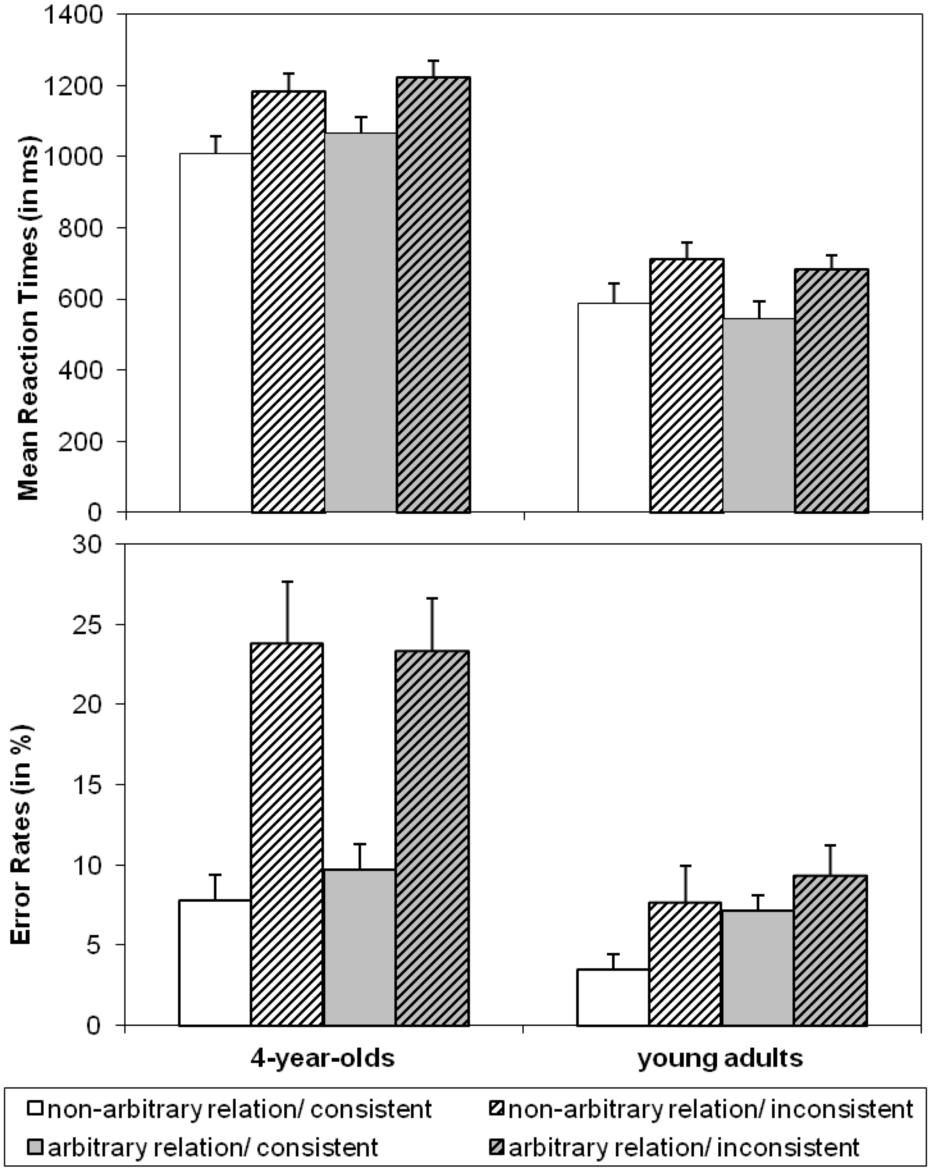

*Figure 1***.** Mean RTs (upper panel) and percentages of errors (lower panel) for performance on arbitrary and non-arbitrary sound-action relations as a function of age and consistency.

Accuracy. Figure 1 shows the main effects for age, with adults making fewer errors than children (7% vs. 16%) (*F*(1, 60) = 18.92, *p* = < .001, η2p = .24). As predicted by our bidirectional-association hypothesis, mapping yielded a main effect with fewer errors made in the consistent mapping (consistent mapping 7%, inconsistent mapping 16%), *F*(1, 60) = 17.87, *p* = < .001, η2p = .23. This effect was involved in a significant interaction with age group, *F*(1, 60) = 7.38, *p* < .01, η2p = .11. Follow-up separate analysis for both age groups revealed that this was due to the mapping effect being significant in children *F*(1, 30) = 15.34, *p* < .001, η2p = .34 (consistent mapping 8%, inconsistent mapping 23%) but not in adults, *p* > .11, a pattern that replicates Eenshuistra et al. (2004). Overall, fewer errors were made with non-arbitrary action-tone relations than with arbitrary tone relations (10 vs. 12% for non-arbitrary and arbitrary relations, respectively), *F*(1,60)= 4.06, *p* < .05, η2p = .06 irrespective of mapping. All other interactions with associatedness failed to reach significance (p>.2). Moreover, children committed more no-go responses than young adults (5.3% vs. 0.5%), *F*(1, 60) = 6.51, *p* < .02, η2p = .98. This effect was not modulated by consistency (*p* > .25).

Discussion

The aim of the present study was threefold. First, we were interested to see whether action-effect relations that are explicitly described as causal, and in a context where the effect is willfully produced by carrying out the respective action, are acquired in the same way as implicit action-effect relations. In particular, we asked whether explicit, intentional acquisition would lead to the creation of unidirectional associations between the representations of the action and the effect. If that would have been the case, no consistency effects would have been expected in the transfer phase. Given that such effects were obtained, and that the size of these effects is almost identical to what has been observed under implicit-learning conditions (Eenshuistra et al., 2004; Elsner & Hommel, 2001), it makes sense to assume that explicit learning does not differ from implicit learning and that both types of learning result in bidirectional associations between action and effect representations. This is an important extension of previous findings and suggests that ideomotor logic applies to both the accidental pickup of possible future action goals during non-intentional movement (as typical for infants and young children) and the intentional acquisition during active exploration of action possibilities and affordances (as more typical for novices facing a new task and more experienced agents fine-tuning their actions). Finding evidence for bidirectional action-effect associations under explicit-learning conditions provides further theoretical constraints on the explanation of Verschoor et al.'s (submitted) observation that asking subjects to repeat the task instructions in their own words prevents transfer from the acquisition to the transfer phase. Even though this manipulation is likely to render the causal relation between actions and effects in the acquisition phase explicit, the present findings do not suggest that this is sufficient to eliminate transfer effects in bidirectional binding.

The second aim of the present study was to compare the learning of novel, arbitrary relations between actions and effects, as often used in ideomotor studies, with the learning (or, better, strengthening) of action-effect relations that fit with pre-experimental knowledge. Although the effects of pre-experimental knowledge on RT's and errors were small, they were coherent and reliable. Non-arbitrary, familiar relations affected performance and allowed for faster responding in both the acquisition and the transfer phase and more accurate responding in the transfer. However, there was no evidence that the type of relationship interacted with any other variable, including consistency. On the one hand, this means that our manipulation was successful in contrasting familiar and unfamiliar relations. On the other hand, however, the familiarity of a relation does not seem to modulate the format of the emerging or strengthened association. In particular, greater familiarity with the relationship between an action and its consequences does not seem to induce a unidirectional action → effect association per se. Given that our subjects were likely to know that trumpets and bells are causally related to the sounds they produce, we need to assume that they have corresponding unidirectional representations in principle but these representations were not activated in the context of the transfer phase. This confirms the consideration of Fenker et al. (2005) and Verschoor et al. (submitted), that unidirectional and bidirectional associations between events can coexist and be selectively activated in some contexts but not in others, presumably depending on instructions and task requirements. From a more general point of view, our findings are also consistent with the observation of Namy, Campbell, and Tomasello (2004) that arbitrary relationships between events are learned just as well as non-arbitrary relationships. In their study, children were exposed to pairings between particular plastic toys and particular gestures, with some gestures bearing an iconic relationship to the toy (e.g., a hopping gesture and a rabbit) and others bearing an arbitrary relationship (a dropping motion and a rabbit). Learning was equally efficient in both conditions in 18-months-olds and 4-year-olds.

Our third aim was to compare transfer effects in young children and adults. Mapping consistency affected performance to the same degree in either group as far as reaction times are concerned. The error rates were also sensitive to consistency but more so in the 4-year-olds than in adults. This profile amounts to a perfect replication of previous observations (Eenshuistra et al., 2004; Kray et al., 2006). It suggests that young children were less able than young adults to keep the instructed stimulus-response mapping active during the transfer phase, presumably due to less mature executive functioning and more resulting “goal neglect” (Eenshuistra et al., 2004). That is, stimuli related to possible action effects seem to reactivate the associated action to such a degree that this tends to overwrite the stimulus-response mapping held in working memory. However, in the trials where children do manage to keep the stimulus-response mapping active and make a correct response, they show a similar effect of mapping on RTs as observed in the adults. In other words, the established task set seems to be equivalent in children and adults but the likelihood of keeping it sufficiently active is higher in adults. Most important for present purposes, however, this profile does not seem to change with explicit action-effect acquisition—as even the present effect magnitudes are comparable to implicit-learning conditions (Eenshuistra et al., 2004)—

or the familiarity of the action-effect relation. Given that adults are likely to have had more experience with trumpets and bells and the sounds they produce than 4-year-old children have, this corroborates the conclusion, contrary to Verschoor et al. (submitted), that familiarity with the (unidirectional) causal connection between actions and effects does not prevent the creation of bidirectional associations between their representations.

To summarize, the present study shows that experiencing sequences of actions and effects induces the creation of bidirectional associations between action and effect representations, no matter what the degree of familiarity with related action-effect contingencies and irrespective of whether the sequence is picked up implicitly or the action is explicitly carried out to produce the effect. This corroborates predictions from ideomotor approaches to action control (Elsner & Hommel, 2001; Hommel, 1996; Hommel et al., 2001; James, 1890) and demonstrates their broad applicability. At the same time, this converging evidence for bidirectional associations in children and adults raises the interesting issue how these associations, and their underlying learning processes, relate to causal learning proper, that is, to the acquisition of what Tolman and Brunswik (1935) have called “the causal structure of the environment. On the one hand, bidirectional associations can be considered to represent empirically derived knowledge about objective action-effect contingencies, which renders it some sort of causal knowledge. On the other hand, however, the bidirectionality of these associations—which the present findings seem to confirm—neglects the empirical and formal difference between cause and effect and, thus, violates the requirements of causal-learning theories (e.g., Gopnik, Glymour, Sobel, Schulz & Kushnir, 2004; Waldmann & Martignon, 1998). Given the observations that perceived causality does not depend on, and is no requirement for action-effect acquisition (Elsner & Hommel, 2004; Verschoor et al., submitted), this suggests that the acquisition of formal causal relationships is independent of ideomotor action-effect acquisition. This would imply that voluntary action control does not rely on a formal understanding of the world's causal structure.

Chapter 6: Where do action goals come from? Evidence for spontaneous action-effect binding in infants

Stephan Verschoor, Maaike Weidema, Szilvia Biro & Bernhard Hommel

Chapter 6

Chapter 6

Abstract

One of the great questions in psychology concerns how we develop to become intentional agents. Ideomotor theory suggests that intentional actions depend on, and emerge from the automatic acquisition of bidirectional action-effect associations: perceiving an action-effect sequence creates an integrated representation that can be employed for action control in the opposite order, selecting an action by anticipating its effect. We provide first evidence for the spontaneous acquisition of bidirectional action-effect associations in 9- 12- and 18-month-olds, suggesting that the mechanism underlying action-effect integration is in place at the latest around 9-months-old.

Introduction

Humans actively manipulate their environment to reach goals, that is, to produce particular intended effects. As the selection of goal-directed actions logically depends on associations between actions and their consequences, memories for actions and their consequences must exist to permit the actor to choose the right action for a given desired outcome. Some authors have suggested that intentional action and the capability to create goal representations are with us from birth on—or even earlier (e.g., Meltzoff & Moore, 1977; Rizzolatti & Craighero, 2004; Rochat, 2001). However, given that a meaningful representation of an action goal requires knowledge about action-effect relations, the complexity and flexibility of human behaviour make it unlikely that action goals are innate (Elsner & Hommel, 2004). Accordingly, other authors have suggested that action goals emerge through experience (Heyes, 2001) and through the observation of the relationship between actions and their perceptual consequences in particular (Greenwald, 1970; Hommel & Elsner, 2009; James, 1890; Lotze, 1852).

One of the first theories addressing how infants (and older learners of novel actions) acquire action-effect associations and how they utilize these associations to produce goal-directed actions was put forward by James (1890). His ideomotor theory states that all actions (or movements) are necessarily involuntary at first and thus no more than motor babbling (cf., Meltzoff & Moore, 1997). However, through experience actions and their effects are automatically integrated and become associated in a bidirectional fashion. This renders action-effect representations mental cues of actions, so that actions can be performed voluntarily by imagining the wanted action effect—which reactivates the respective effect representation and primes the associated action. Ideomotor theory was revived and refined by Greenwald (1970), Prinz (1990, 1997), and Hommel (1996; Elsner & Hommel, 2001), and then elaborated into the Theory of Event Coding (TEC: Hommel et al., 2001). TEC holds that action and perception share a common representational domain and stresses the interplay between, and mutual dependency of perception and action. This provides a broad theoretical background and motivation for the assumption that actions and codes of their perceptual consequences are associated in a bidirectional fashion. This bidirectionality is particularly important for ideomotor theorizing because it assumes that perceiving the sequence of action and effect creates a representational structure that can be used for action control in the opposite order (anticipation of action effect leading to action). TEC also suggests that the same representational format is used for observed and self-initiated action thus also providing a framework for action understanding.

Evidence for the acquisition of bidirectional associations between actions and their effects has been found in different species, such as cats, rats and pigeons, in humans ranging in age from 4-year-olds (Eenshuistra et al., 2004; Kray et al., 2006) to adults (Elsner & Hommel, 2001), and for a wide range of actions and effects (for a review see: Hommel & Elsner, 2009). Action-effect acquisition is modulated by the same factors that influence instrumental learning (e.g., temporal contiguity and contingency of movement and effect: Elsner & Hommel, 2004) and does not depend on voluntary attention (Dutzi & Hommel, 2009; Elsner & Hommel, 2001; Band et al., 2009), suggesting that it is a fairly automatic process indeed (Elsner & Hommel, 2004). Additionally, action-effect acquisition has been found after just one trial (Dutzi & Hommel, 2009).

Given this apparently high degree of automaticity and the strong evidence that very young infants are sensitive to action-effect contingencies (e.g., Rochat & Striano, 1999; Gergely & Watson, 1999; DeCasper & Fifer, 1980; for a review, see Rovee-Collier, 1987), it seems likely that the proposed mechanism is already operative in early infancy. However, even though the role of action-effect learning in the development of voluntary action is receiving increasing attention, this interest is mostly focused on infants’ interpretation and imitation of other people’s goal-directed actions. The central assumption underlying various studies is that acquired action-effect associations are instrumental in understanding and imitation of actions. Indeed, action effects have been shown to be important for action understanding (e.g.; Biro & Leslie, 2007; for a review, see: Hauf, 2007; Kiraly et al., 2003) and in imitation behavior, as infants are more likely to reproduce actions that have a salient action effect (Hauf & Aschersleben, 2008; Klein et al., 2006; for a review, see: Elsner 2007; Melzoff, 2007). Although these studies suggest that observed action-effect relations in other people's actions can influence infant's imitation behavior and action understanding, they do not tell whether infants learn action-effect associations by exploring the world themselves. Furthermore, these studies do not provide direct evidence for the assumed bidirectionality of action-effect associations, which according to ideomotor theory allows the infant to select and produce actions by activating their perceptual consequences.

The present study sought for more direct evidence for the acquisition of bidirectional action-effect in infancy. Although we hypothesize that the ability to learn action-effect associations is present from early infancy, action production and action control are extremely limited in very young babies, thus rendering behavioural measurements difficult. Since investigating bidirectional action-effect acquisition requires participants to carry out some sort of motor actions, a possible lack of evidence in very young infants could simply be due to limitations in their motor capabilities rather than to the inability to associate and rely on action-effect relations. Adult and children studies on learning action-effect associations typically used manual actions such as pressing a button. Good control over these types of manual actions develops in the second half of the first year (Belsky & Most, 1980). In addition, previous research showed that the earliest age at which infants' own actions are influenced by observed effects of others' actions is around 9 months (Hauf & Aschersleben, 2008). At this same age infants begin to distinguish means from ends in their own behaviour (Piaget, 1936; Goubet, Rochat, Maire-Leblond, & Poss, 2006; Willatts, 1999). This ability is often thought to be a prerequisite for goal-directed action control, as it enables infants to specify and represent goals before performing the corresponding actions

(Hauf, 2007). Therefore, we expected that at 9 months infants will likely show evidence for bidirectional action-effect acquisition via their own manual exploratory behaviour. Furthermore, Elsner and Aschersleben's study (2003) suggests that there are developmental changes in the effect that observed action-effect relations have on infants' own behavior. They found that with a higher task demand the behavior of 12 but not 9 months old infants was influenced by an observed action-effect. In addition, a further step was made at around 15-18 months when infants' behavior started to be affected by whether the effect they could produce matched with the observed effect. Testing 9, 12 and 18 months olds' ability to acquire bidirectional action-effects by their own exploratory behavior can help clarify whether the nature of these changes lies in the development of specific imitative skills for transferring others action into one's own action or in the accuracy of encoding specific action-effect relations due to, for example, increasing working memory capacity or brain maturation (Diamond, 2006).

Our task was modelled after the experimental setup used in the free-choice experiments of Elsner and Hommel (2001). Elsner and Hommel had adults carry out self-chosen left and right key presses in response to a visual trigger stimulus. Each key press produced a particular sound (e.g., left key → low tone, right key → high tone), and even though these sounds were irrelevant for the task, it was assumed that participants would acquire bidirectional associations between key presses and tone representations. After this acquisition phase, participants were again freely choosing left and right key presses, but in this test phase the visual trigger was replaced by an auditory trigger stimulus: high and low tones (identical to the previous action effects) presented in random sequence. As predicted, people were quicker and more likely to choose the action that previously had produced the currently presented trigger tone (e.g., they were quicker and more likely to press the left key when hearing the low than the high tone), suggesting that key presses and tones were indeed associated in a bidirectional fashion.

When adopting this paradigm for our infant subjects, we realized after piloting that versions of the original binary-choice task were too demanding. We therefore simplified the task in such a way that only one (very large) touch-sensitive key was presented to the infants, the two action alternatives being the touching or not touching of the key. Moreover, given Dutzi and Hommel's (2009) demonstration of action-effect learning after just one presentation, we also greatly reduced the number of acquisition- and test trials, so to minimize demands on infants' limited attention. In the acquisition phase, infants were presented with two multimodal (and thus very salient and attention-grabbing) events, one being self-produced by touching the key (a true action effect) and another that was presented while infants were prevented from touching the key, so that the event was not self-produced. In the test phase, we presented the two multimodal events in different blocks and expected the previously self-produced event stimulating quicker and more spontaneous key touches than the action-independent event.

Materials and methods

Participants

Three groups of infants were tested: 22 9-month-olds (mean: 9.09 months, SD= .18, 7 female), 21 12-month-olds (mean: 12.10 months, SD= .26, 10 female), and 22 18-month-olds (mean: 18.02 months, SD= .32, 8 female). Four additional 9-month-olds, three 12-month-olds, and two 18-month-olds were excluded due to fussiness. One 12-month-old was excluded due to experimental error. The participants were recruited through advertisements in local papers, daycares, maternity wards, and general practitioners. Infants were randomly and equally distributed across conditions, and they received small gifts as reward. An informed consent was obtained from all caretakers.

Apparatus

During the experiment infants sat in their caretaker's lap in a curtained booth. The multimodal stimuli (sounds and images) were presented via a 30-inch widescreen monitor with built-in speakers situated at a distance of 60 cm from the infants. A brightly-colored touch-sensitive key was placed right in front of the infant. The key was built into a wooden board that was attached to the booth below the monitor (see Figure 1). Presentation and latency recording was controlled by E-prime™ software. A camera located above the booth recorded the touch-sensitive key. Another camera behind the infant recorded the stimulus presentation. The cameras fed a mixer that produced a split-image movie of the infants’ actions on the key and stimuli.

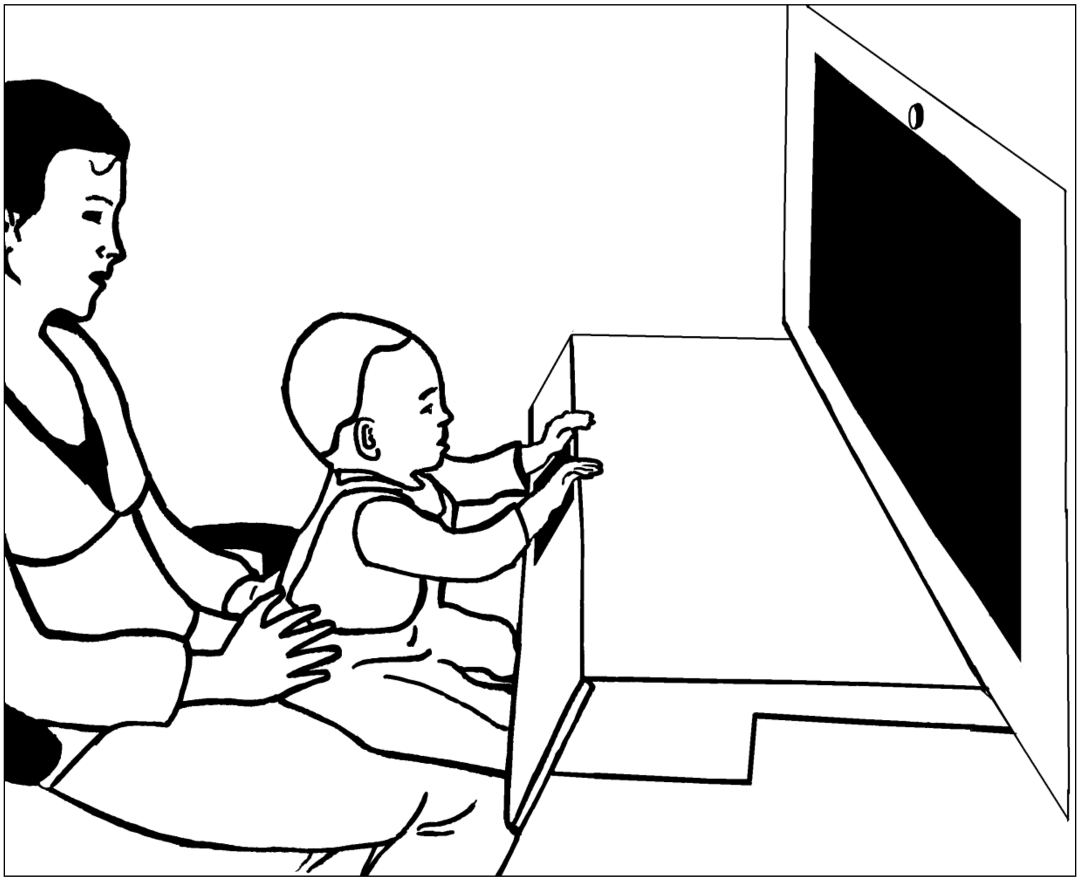

Figure 1. Experimental set-up. The infant sat on the caretaker's lap while perceiving the stimulus events and pushing the touch-sensitive button that produces the action effect.

Procedure

Instructions were given to the caretakers prior to the experiment. The experiment consisted of two phases: the acquisition phase and the test phase. The acquisition phase was composed of two blocks of five trials with self-produced multimodal events and two blocks of five trials with action-independent multimodal events, in an alternating order. This setup resulted for all infants in 20 acquisition trials, 10 self-produced effect trials and 10 action-independent effect presentations. We chose multimodal events to maximize their discriminability and attention-grabbing potential.

In blocks with action-independent events, infants were presented with one of the two audiovisual events five times in a row, while the caregiver prevented the infant from touching the key by gently holding the infant's hands. To ensure the infant paid attention to the screen the presentation was triggered by the experimenter, who monitored the infant on-line. In case of distraction the experimenter waited for attention to return. In blocks with self-produced events, the infant's hands were free and could press the touch-sensitive key. Each touch of the key immediately elicited a presentation of an audiovisual event. Before another effect could be elicited the previous effect had to end. The caretakers were instructed to encourage the infants to press the key. Between the blocks there was a short break of 10 seconds to ensure that infant and caretaker were ready for the next block. Additionally, a red or green dot was presented on the monitor during the blocks to remind the caretakers which block they were in. The bimodal effects were two distinct 500-ms long meaningless sounds that started simultaneously with the 1000-ms presentations of a picture (either bright-colored cartoon of a car or a mouse)[12].

---

[12] We decided not to balance action-independent and self-produced effects across participants after a pilot experiment had revealed no preference for one or the other effect. The pilot investigated 15 participants, five from each of the three age groups. Infants were shown the button but prevented from touching; and it was checked whether the infant was actually looking at the button. We then presented the two bimodal events in random order, and the infant had the opportunity to touch the button for 30 seconds. A short break, during which infants were prevented from reacting again, divided the two presentations. The two effect stimuli did not yield any reliable difference in the latencies ($F(1,12)=.627$,

After the acquisition phase there was a 30-second break during which caretaker and infant turned away from the monitor while engaging in entertaining the infant. The test phase consisted of three trials, in which the previously self-produced event, the previously experienced action-independent event, or no event was presented. Each of these trials lasted 30 seconds, during which the infant could freely touch the key. In between the test trials there was a 10-sec break during which the caretaker prevented the infant from responding by holding his/her hands. The self-produced and action-independent events were presented first and second, or vice versa, with the order being balanced across participants, while the baseline trial was always administered last—so to minimize possible forgetting and extinction of action-effect associations. Methods were approved by the ethical committee of the Leiden University, Institute for Psychological Research.

Data acquisition

In acquisition trials, we measured latencies online and, based on the offline inspection of the video tape, the number of undetected motor responses (visible scratches and touches of the key that were too light for the conductance-sensitive key to detect) and the number of responses on which the infants were helped by the caretaker. Given that the acquisition of action-effect associations is sensitive to contingency and extinction (Elsner & Hommel, 2004), the undetected (and therefore not "effect-rewarded") motor responses were considered particularly important. In test trials, we measured the latency of the first response and the number of responses (key touches). From the latter, response

---

p=.444) or the number of responses (F(1,13)=.398, p=.389), nor was there any hint to interactions with age or order of presentation.

frequency was calculated for each infant based on his/her total number of responses in all three test trials, thus correcting for motivational differences (see Table 1).

Table 1. Mean number of responses (#) and mean percentage (%) of responses for 9-month-olds (n = 22), 12-month-olds (n = 21) and 18-month-olds (n = 22) in Self-Produced, Action-Independent and No-Event test-trials.

Test-Trial Type	Self Produced		Action Independent		No Event	
Age groups	#	%	#	%	#	%
9-Month olds	5.00	38.95	4.91	41.23	2.50	19.73
12-Month olds	3.67	36.38	4.33	41.60	2.24	22.02
18-Month olds	4.86	49.53	4.36	26.67	2.41	23.75

Results

Acquisition phase

One-way between-subjects ANOVAs were carried out to assess possible differences between the three age groups in the acquisition phase. No reliable differences were obtained for the mean latencies for responses in trials with self-produced events, in the percentage of responses in which the infants were helped by the caretaker, or in the timing of the presentation of the action-independent events (Fs<1). However, we did find an effect of age group on the number of undetected key touches, $F(2,62)=3.23$, $p = .032$, $\eta^2_p = .105$: Pairwise comparisons revealed that the 12-month-olds overall had significantly more undetected responses (3.33) than the 9-month-olds (0.95), $p=.01$, but not than the 18-month-olds (1.82).

Test phase

To assess possible differences in motivation and/or motor abilities, we ran an ANOVA on the total number of responses made in all three test trials, but no effect of age was obtained. The mean number of responses during the test-phase was 11.45.

A between-subjects ANCOVA was carried out on the latencies of first responses in the test trials, with trial type (previously self-produced vs. action-independent stimulus event) as within-subjects and age group (9, 12, 18 months) as between-subjects factor (see Figure 2). In this analysis we considered the number of undetected responses during acquisition as a covariate, so to control for extinction, and the total number of responses during all three test trials, so to control for the overall level of activity. We found a significant main effect of trial type, $F(1,51)=4.67$, $p = .035$, $\eta^2_p = .564$: Infants responded significantly faster to self-produced events (5.57 s) than to action-independent events (6.64 s). No effects of age group or interactions with the covariates were found.

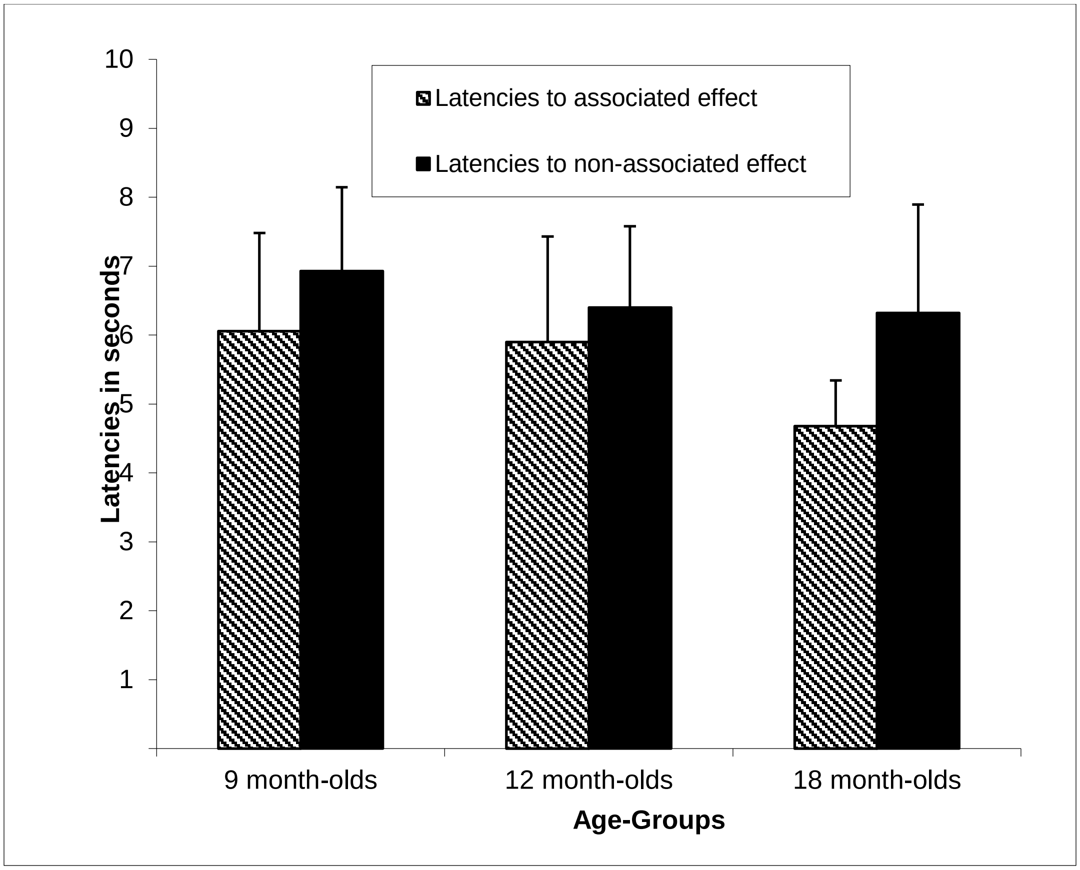

Figure 2. Mean Latencies (+ SE) for 9-month-olds (n = 22), 12-month-olds (n = 21) and 18-month-olds (n = 22) in associated and non-associated test-trials.

We ran another mixed-factor ANCOVA on the response proportions in the test trials, normalized through an Arcsin transformation. Trial type and age group served as factors (see Figure 3) and the number of undetected responses during the acquisition as covariate. There was no main effect of trial type (*F*(1,61)=0.70, p = .41, η2p = .011) but a close-to-significant interaction of trial type and age group, *F*(2,61) = 2.89, p = .063, η2p = .087. Separate exploratory between-subjects ANCOVAs revealed that there was no trial-type effect in the 9- or the 12-month-olds, *F*s<1, but responses were more frequent to self-produced than action-independent events in the 18-month-olds, *F*(1,20) = 5.31, p = .032, η2p = .210.

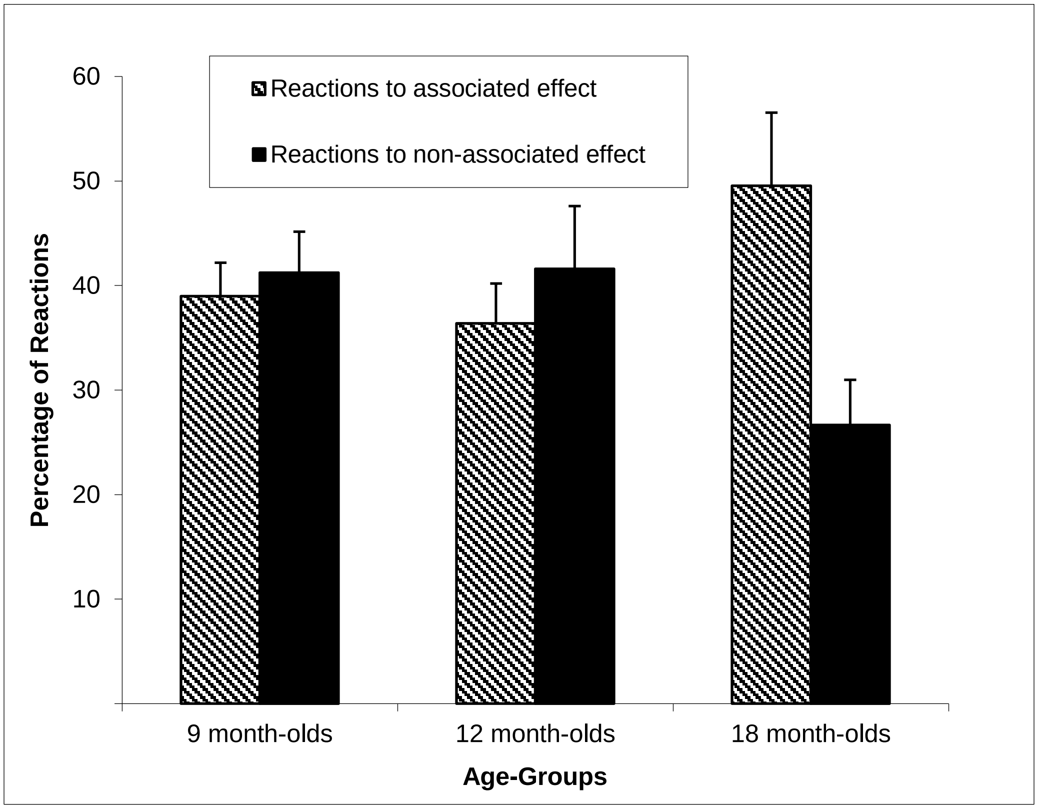

Figure 3. Mean Percentage of responses (+ SE) for 9-month-olds (n = 22), 12-month-olds (n = 21) and 18-month-olds (n = 22) in associated and non-associated test-trials.

Chapter 6

Discussion

In the present study we sought evidence for the spontaneous acquisition of bidirectional action-effect associations in early infancy. As expected, 9-, 12- and 18- months-olds were faster to respond to events that they previously had actively produced than to action-independent events, indicating that all age groups indeed formed bidirectional action-effect associations during the acquisition phase. Moreover, at least the 18-months-olds also had a stronger tendency to perform the action again and more often compared when presented with the effect they previously caused than with the action-independent event. Altogether, we consider this pattern a rather close replication of the Elsner & Hommel (2001) finding in adults.

Unexpectedly, the latency measure turned out to be more sensitive than the response-frequency measure, which showed evidence for action-effect acquisition in the 18-month-olds only. This differential sensitivity might be due to several, not necessarily mutually exclusive factors. For one, it is possible that the younger infants were creating weaker or less specific associations between action and effect representations in the acquisition phase. These weaker associations might have been sufficient to drive the first response, which therefore was faster to self-produced effects, but might have fallen prey to extinction too soon to produce a larger number of responses to self-produced effects than to action-independent effects. Indeed, some evidence for extinction has been obtained even in adults (Elsner & Hommel, 2001). For another, it is possible that younger infants were less likely to retrieve the required action-effect associations in the test phase—be it due to less flexible transfer from the acquisition to the test context or because of higher cortical noise (note that the greatest changes in myelination occur during the first three years: Carmody, Dunn, Boddie-Willis, DeMarco & Lewis, 2004). In any case, the fact that only 18 months old infants showed evidence for action-effect acquisition in response-frequency measure, fits with the developmental changes found in the Elsner and Aschersleben's study (2003) that we mentioned before. Similar to the current finding, in their study the age groups did not differ from each other in their latency measures while only the 18-month-olds performed more target actions when the effect they could produce matched the effect of the observed model's action. Although the two studies are different in their design and in their research questions, our finding suggests that the differential frequency pattern found between age groups in their study was not, or not only, due to the development of imitative skills, but may have been due to differences in the strength of specific action-effect associations. Nevertheless, in our study the latency measures showed that all age groups acquired a bidirectional action-effect association.

A disadvantage of our simplified design is that it does not allow distinguishing associations between the effect stimuli and specific motor responses from associations between effect stimuli and a broader range of motor activities. For instance, it is possible that the self-produced effect became associated not so much with particular manual key-reaching actions but with general motor activity or playing in general, while the action-independent effect became associated with lack of activity. On the one hand, such an approach does not seem to fit with the behavior of our participants in the action-independent effect condition—as evident from the session tapes, which did produce motor activity that however was not specifically directed to the key. On the other hand, even such a rather unspecific association between effect stimuli and motor activation as such must rely on bidirectional associations between action codes and action-effects codes, as predicted from ideomotor theory.

In summary, our findings demonstrate the spontaneous, non-intentional acquisition of bidirectional action-effect associations in infants no older than nine months. This observation by no means contradicts nativist ideas about action-goal representations but it does provide a theoretical alternative that makes the less parsimonious nativist assumptions unnecessary. Apparently, infants are able to pick up action effects that they are able to control on the fly and establish bidirectional associations between representations of these action effects and the motor actions (or class of motor action) producing them. This way, infants acquire not only possible future action goals but also the means to reach them whenever they might become interested in doing so later on. In other words, human action goals might be grounded in and through the acquisition and anticipation of action-contingent perceptual effects. This fits with findings from studies on action perception, showing that infants around 9 months more readily encode actions that have salient action effects as goal-directed (Biro et al., under revision; Biro & Leslie, 2007; Hauf, 2007), and with studies on imitation, showing that 9-months-olds can use observed action effects for action control (Elsner, 2007). Taken altogether, these previous and the present findings provide considerable support for the basic assumption of TEC (Hommel et al., 2001) that the integration of actions and their effects provides a common representational format for the coding of observed and self-produced actions, thus making interactions between them possible.

Chapter 7: From outcome prediction to action selection:

Developmental change in the role of action-effect bindings

Stephan A. Verschoor, Michiel Spapé, Szilvia Biro & Bernhard Hommel

Abstract

Ideomotor theory considers bidirectional action-effect associations to be the fundamental building blocks for intentional action. The present study employed a novel pupillometric and oculomotor paradigm to study developmental changes in the role of action-effects in the acquisition of voluntary action. Our findings suggest that both 7- and 12-months olds (and adults) can use acquired action-effect bindings to predict action outcomes but only 12-months-olds (and adults) showed evidence for employing action-effects to select actions. This dissociation supports the idea that infants acquire action-effect knowledge before they have developed the cognitive machinery necessary to make use of that knowledge to perform intentional actions.

Chapter 7

Chapter 7

Introduction

Some authors have suggested a rudimentary ability to represent action goals to be present at birth (e.g., Meltzoff & Moore, 1997; Rizzolatti & Craighero, 2004; Rochat, 2001). But where do such representations come from? Given the world's complexity and the dramatic changes the mind and body of infants undergo during development, it is rather unlikely that they are innate and permanent (Greenwald, 1970; Harless, 1861; Hommel & Elsner, 2009; James, 1890; Lotze, 1852). Piaget (1936), in his influential constructivist approach to cognitive development, firstly suggested that goals should be adaptive to the infant's changing skills and abilities and may derive from its own sensorimotor exploration and experience motivated by a predisposition to adjust to its environment. Although he states that such exploratory actions are integrated with their effects into schemata necessary for perception and action, he does not elaborate on the underlying cognitive mechanism (Piaget, 1954).

Exactly such a mechanism was proposed in the late nineteenth century, though largely neglected in psychology for a century. Following the lead of Lotze (1852) and Harless (1861), James (1890) suggested a cognitive mechanism that does what Piaget (1936) proposed; it provides actors with action goals that are rooted in their own sensorimotor experience. In his ideomotor theory James stated that all actions are necessarily involuntary when being carried out for the first time. Indeed, if one defines action as goal-directed movement, it presupposes some sort of anticipation of its effect. This again implies knowledge on action-effect relationships, which needs to be acquired before the action can be carried out "in order to" produce the outcome intentionally. Ideomotor theory suggests that such knowledge is acquired on the fly: whenever people move, they automatically and unintentionally create bidirectional associations between the perceived effects and the motor pattern producing them. This association brings the movement under voluntary control: Once acquired, the agent can now activate the motor pattern producing a movement by "thinking of" (i.e., endogenously activating the representation of) a perceptual effect. Indeed, infants start to motor babble (i.e., produce random movements) *in utero* (cf., Meltzoff & Moore, 1997)—which could explain the possible presence of goal representations at birth—and they are consistently exploring their environment. This provides ample opportunity to acquire movement/action-effect associations and thus a steadily increasing pool of possible action goals. Thus, James considered bidirectional movement/action-effect associations the fundamental building blocks of intentional action and provides a mechanism that could allow the emergence of goal-directed action in infants.

Ideomotor theory was revived and refined by Greenwald (1970), Prinz (1990, 1997), and Hommel (1996; Elsner & Hommel, 2001) and is now part of a broader theoretical movement stressing the interplay between perception and action (Hommel et al., 2001; Meltzoff, 2007; Meltzoff & Prinz, 2002). This motivated numerous demonstrations of bidirectional action-effect acquisition in humans ranging from 4-year-olds (Eenshuistra et al., 2004; Kray et al., 2006) to adults (Elsner & Hommel, 2001). Action-effect acquisition was found for a wide range of actions and effects (for a review see: Hommel & Elsner, 2009), suggesting a general action-effect integration mechanism. Additionally, action-effect acquisition has been found after just one trial (Dutzi & Hommel, 2009), suggesting that the mechanism is fast-acting and implicit. Action-effect acquisition is modulated by the same factors that influence instrumental learning (e.g., temporal contiguity and contingency of movement and effect: Elsner & Hommel, 2004) and does not depend on voluntary attention (Dutzi & Hommel, 2009; Elsner & Hommel, 2001; Band et al., 2009). Together with the fact that it was also found in animals (see Elsner & Hommel, 2001), this suggests that action-effect integration it is a fairly low-level and automatic process (Elsner & Hommel, 2004).

If bidirectional action-effect associations are indeed the fundamental building blocks for intentional action, the system that generates these associations should be operative early in life. Especially since infants show evidence of goal-directed behavior from a very early age on: depending on the definition, goal-directed behavior is thought to start somewhere between birth (Metzoff & Moore, 1997; von Hofsten, 2004) and about nine months of age (Hauf, 2007; Piaget, 1936). Action-effect knowledge has been implicated to be operational in higher order cognitive functions such as action understanding in 7 months-olds (e.g.; Biro & Leslie, 2007; for a review, see: Hauf, 2007; Kiraly et al., 2003) and imitation in 9-months-olds (Hauf & Aschersleben, 2008; Klein et al., 2006; for a review, see: Elsner 2007; Meltzoff, 2007). Even though these findings do not provide direct evidence for bidirectional action-effect acquisition, theories that emphasize similar representational formats for first-person experience and observed action (e.g. Fabbri-Destro & Rizzolatti, 2008; Hommel et al.,2001; Meltzoff, 2007; Tomasello, 1999), and conceptualize action understanding as inverse planning (Melzoff, 2007; Baker et al., 2009) consider them corroborative. Other corroborating evidence was found in studies that show very young infants to be sensitive to action-effect contingencies. For instance, newborns actively adjust their sucking rate in response to their mothers’ voice as ongoing conditional feedback (DeCasper & Fifer, 1980) and 2-month-olds pursue interesting action effects by intentionally varying their sucking rate (Rochat & Striano, 1999) or varying gaze direction (Watson, 1967; for a review, see Gergely & Watson, 1999). Another line of research by Carolyn Rovee-Collier shows that action effects aid memory retrieval for actions from two months of age (Rovee & Rovee, 1969; for a review, see Rovee-Collier, 1999). Telling as these studies may be (they show that action contingent effects play an important role in infant behaviour and memory) they were not designed to directly asses the bidirectionality of action-effect associations and their use for action planning and may thus confound actual action-effect learning with simple operant conditioning. Nonetheless, these findings may reflect a developing ability for learning action-effect contingencies.

The first study to find direct evidence for bidirectional action-effect associations in infants was undertaken by Verschoor et al. (2010) with 9-, 12- and 18-months-old participants. A simplified version of the free-choice design of Elsner & Hommel (2001) was used. In the acquisition there were two types of trials: One in which infants were permitted to touch a response key, which resulted in a particular audiovisual effect, and another in which they were prevented from touching the key while another audiovisual effect was presented. When the two effects were replayed in the test phase, all infants were faster to touch the key again after the previously self-produced effect than after the action-independent effect. Moreover, the 18- months-olds responded more often following the self-produced effect. These results were taken as evidence for bidirectional action-effect learning, since the self-produced effect activated the response that previously caused the effect. Although this study shows direct evidence for action-effect learning in infants, the paradigm had some drawbacks compared to the Elsner and Hommel (2001) paradigm. The paradigm had only 10 acquisition trials and 2 test trials which resulted in slightly noisy data, and due to the nature of the task the paradigm was unsuited for comparison with adults. Thus it remains to be seen whether bidirectional action-effect learning works similarly in infants and adults. Furthermore, initial piloting showed that, due to difficulties with the button pushing action, it was unsuitable for infants younger than nine months (Verschoor et al., 2010).

Overcoming these limitations, calls for a more natural type of action that is well established in very young infants. Eye movements seem to be the ideal candidate: Infants have been reported to actively and accurately control their eye movements from at least 4 months of age (Scerif et al., 2005) and, given that infants actively control their gaze to gather information (Gredebäck & Melinder, 2010; Falck-Ytter, Gredebäck, & von Hofsten, 2006), to direct or follow attention (Perra & Gattis, 2010), and to engage in social behaviors (Senju & Csibra, 2008; Johnson, Ok & Luo, 2007b), eye movements can be considered truly goal-directed actions. Moreover, a study by Herwig and Horstmann (2011) demonstrated saccade-effect learning in adults in a paradigm conceptually very close to that of Elsner and Hommel (2001), which indicates that action-effect integration generalizes to oculomotor action.

In the present study, participants made eye-movements towards visual stimuli appearing at the left or right of a display, the two directions produced different auditory effects—analogous to the Elsner and Hommel (2001) paradigm. In the following test phase, participants were presented with the effect tones and could freely choose to make a saccade to the left or right of two simultaneously presented visual stimuli. We expected that the tone would prime the saccade that it was produced by in the acquisition phase, so that this saccade would be chosen more frequently and/or initiated quicker. This design allowed us to test both infants younger than 9 months of age and adults, and to run considerably more trials.

We tested 7- and 12- months-olds and adults. 7-months-olds were chosen because this group is known to show understanding of goal-directed actions (e.g. Woodward 1998; Csibra, 2008, Verschoor & Biro, 2012). Since some action-perception theorists (Hommel et al., 2001; Meltzoff, 2007; Meltzoff & Prinz, 2002; Woodward, 2009; Rizolatti & Craighero, 2004) stress that the same representational format is used for observed and self-initiated action, even 7-months-olds should be able to pick up action-effect associations. A first study that shows action-effect acquisition in infants younger than 9 months was recently published by Paulus, Hinnius, Elk and Beckering (2012). They found electrophysiological evidence indicating that infants at 8 months of age, show stronger motor resonance when listening to previously self-produced action-related sounds than when hearing other sounds. It remains to be seen whether the underlying action-effect associations are bidirectional in the sense that they can be reversed to generate overt action. Dissociations between acquired action knowledge and the use of such knowledge are by no means new; for instance Keen (2005) found a similar dissociation in her work on reaching and looking for occluded objects which shows that infants' looking behavior exhibits knowledge while actions are not conform to this knowledge. In a more general sense, such dissociations are apparent in looking-time studies wherein infants are reported to possess knowledge on actions they cannot perform themselves (e.g. Verschoor & Biro, 2012; Csibra et al., 1999).

To maximize the chance of finding developmental changes in this study, we included the older of the two youngest groups in Verschoor et al. (2010) since no differences were found between 9- and 12-

months olds in that study. They were also included to replicate Verschoor et al.'s (2010) finding that infants at this age show bidirectional action-effect learning and to contrast this ability with the suggested inability of the 7-months-olds to initiate true intentional action (Hauf, 2007). We included adults to confirm that the same principles of bidirectional action-effect learning can be shown in 12-months-old infants and adults.

Using an eye-tracker enabled us to measure not only the choice of actions and the time to initiate them (reaction time), but pupil size as well. This is a relatively new measure in developmental studies (e.g. Falck-Ytter, 2008; Jackson & Sirois, 2009; Gredebäck & Melinder, 2010, for a review, see Laeng, Sirois & Gredeback, 2012) but has been extensively used in psychological research on adults since the early 20th century (Hess, 1975). Pupils have the interesting characteristic to react not only to luminance, they reliably dilate with superimposed sympathetic activation (Libby, Lacey & Lacey, 1973; Beatty & Lucero-Wagoner, 2000). Although these dilations are not directly causally related to central processing load, they empirically reflect variations in central processing load with extraordinary precision (Beatty & Lucero-Wagoner, 2000). Task-Evoked Pupillary Responses (TEPRs) can indicate motivational phenomena such as arousal (Bradley, Miccoli, Escrig & Lang, 2008; Laeng & Falkenberg, 2007), attention allocation (e.g. Hess & Polt, 1960), cognitive load (Kahneman & Beatty, 1966), and mental effort (Kahneman, 1973; Hess & Polt, 1964). TEPRs are pre-conscious and mediated by the locus coeruleus (Laeng et al., 2012). Whatever the exact interpretation of this measure, using it enables us to contrast acquisition contingent vs. non-contingent responses: whether TEPRs are taken to reflect differences in sympathetic activation in general or arousal, attention allocation, cognitive load or mental effort, all interpretations suggest that dilations should be larger for incongruent responses.

We assumed that in our task two types of processes might play a role and affect TEPRs. For one, even in studies that were successful in demonstrating that action effects bias the choice of actions (e.g., Elsner & Hommel, 2001), participants did not always choose the action that was previously associated with the present trigger stimulus. To some degree, this might be due to random noise but it may also reflect a strategy to create some variability and exert active control. Exerting control calls for the investment of more cognitive resources, which would suggest that selecting and/or performing an action that is not associated with, and thus primed by the trigger stimulus is more effortful. If so, one would expect that choices of tone-incongruent actions (i.e., of another action than the one that previously produced the present trigger stimulus) are accompanied by (greater) pupil dilation.

For another, there is evidence that action-effect associations not only affect the choice of actions but also their evaluation. In the study of Band et al. (2009), participants performed a probabilistic learning task, in which key-presses triggered tones of a particular pitch in 80% of the trials and of another pitch in the remaining trials. The presentation of a less frequent action effect generated an electrophysiological component that is known as feedback-related negativity (Miltner, Braun & Coles, 1997), which is commonly observed when negative feedback is presented. This suggests that action-effect associations are used to generate particular expectations about effects given the execution of a particular action. In infant studies, TEPRs have been used as an index of the violation of expectations (Jackson & Sirois, 2009; Gredebäck & Melinder, 2010). Accordingly, it is possible that carrying out a tone-incongruent action results in (more) pupil dilation reflecting the violation of a tone-induced expectation regarding the action outcome (i.e., the location of the action end point and/or the targeted stimulus).

Although both processes would predict greater pupil dilations in incongruent responses, we considered that these might be distinguished in terms of their temporal dynamics: whereas a choice-related process would be likely to affect pupil responses briefly before or after response execution, an expectation/evaluation-related process would be more likely to affect pupil responses after response execution (Band et al., 2009).

Methods

Subjects

Two groups of infants were tested: 15 7-month-olds (mean: 7.15 months, *SD*= .21, *SE*=.05, 8 female) and 20 12-month-olds (mean: 12.11 months, *SD*= .26, *SE*=.05, 9 female). The infants were recruited through direct mail. An informed consent and a questionnaire regarding their general health were obtained from all caretakers. The infants were all healthy full-term and without any pre- or perinatal complications. Additionally, 24 undergraduate students (mean age: 23.8 years, *SD*= 2.47, *SE*= .50, 14 female) participated in exchange for course credits. All reported to be healthy and to have normal or corrected-to-normal vision and hearing. Two additional 12-month-olds and one adult were excluded due to technical error, and two more 7- and four 12-month-olds were excluded due to fussiness. Additionally two 7- and two 12-month-olds were excluded for not meeting the criterion for the minimal amount of test trials.

Test environment and apparatus

During the experiment participants sat in a specially designed stimulus-poor curtained booth (infants in the lap of their caretaker) in front of the monitor/eye-tracker apparatus. The distance between eyes and apparatus was approximately 70 centimeters (the screen's viewing angle was 34.1° by 21.8°). Participant behavior was monitored online by means of a camera located above the apparatus. The experimenter controlled the experiment from a separate control room. A 17 inch TFT-screen, equipped with an integrated Tobii T120 eye-tracker operating at 60 Hz, was used for visual and auditory data presentation, and for data collection. The Tobii T120 has an average accuracy of .5 visual degrees and allows for a reasonable amount of free head movement by the subject (30x22x30cm). It recorded gaze direction and pupil-size. Stimulus presentation was controlled by a PC running E-prime® software (Schneider, Eschman & Zuccolotto, 2002).

Procedure

Infants were tested at a time of day when they were likely to be alert and in good mood. Caretakers and participants were given instructions prior to the experiment. Adults were given no instruction with regard to the task. The caretakers were instructed not to move after calibration and gently fixate the infant against their tummy to maintain the eye-tracker alignment and to entertain the infant during the 1-min interruption between calibration and the experiment. The eye-tracker was calibrated using a 9-point calibration consisting of an animated dancing infant accompanied by music. The calibration was accepted with a minimum of eight points acquired successfully. The experimenter could play an attention-grabbing sound during the experiment to regain attention. If the attention grabbing sound did not work caretakers were encouraged to direct the infant's attention to the middle of the screen by pointing to it. Lighting conditions were kept constant during testing and across subjects. Furthermore the luminance levels were controlled for by presenting the stimuli in a random fashion. After completion of the experiment, further information on the rationale was provided.

Acquisition phase

The experiment began with an acquisition-phase of 48 trials (see Figure 1). The background color of the screen was grey. An acquisition trial started with a brightly colored dot with a superimposed line drawing (4.3° by 4.3°) being displayed at the center of the screen (Snodgrass & Vanderwart, 1980). The dot served as start signal and fixation mark. To keep the display interesting to the subjects, the color of the dot changed randomly from trial to trial (selected from eight bright colors) and the superimposed line drawing was randomly selected (without replacement) from a selection of 50 drawings. The dot disappeared after the subject fixated properly for an interval that varied from trial to trial (so to remove any bias or habituation that might be caused by fixed intervals between trials) between 150 and 350 ms.

Immediately after the dot disappeared, photographs of two different faces (randomly selected without replacement from 100 grayscale pictures from the “Nottingham scans” emotional faces database, http://pics.psych.stir.ac.uk, displaying emotionally neutral faces of 50 men and 50 women from a frontal perspective), appeared left and right from the dot. Faces were chosen to elicit spontaneous saccades as they are known to attract infants’ attention (Goren, Sarty & Wu, 1975; Johnson, Dziurawiec, Ellis & Morton, 1991). The 5.3° by 5.3° pictures appeared at 9.7°, center to center, to the left and right of the center of the screen. To avoid perseverance to either left or right across acquisition trials, the images immediately started to pulsate. One of the faces started shrinking to 4.1° while the other started growing to 6.5° (which picture started shrinking was randomized); one cycle from intermediate size to small, to intermediate, to large and back to intermediate, took 2 s.

The faces evoked spontaneous saccades and thus served as response locations. When a saccade towards one of the two face locations was detected, the face at the other location disappeared. The targeted face stopped pulsating and, depending on the targeted side, one of two distinct 200 ms effect-sounds (“tring” or “piew”) was presented. Each effect-sound was consistently designated to either the left or the right response area (RA) during the entire acquisition phase (the mapping was balanced across participants); RA’s were defined as the maximum size of the pulsating images: 6.5° by 6.5°. A saccadic response was defined as an eye movement into the left or right response area, (minimal amplitude 4.3°). Reaction Times (RTs) were defined as the time it took from the disappearance of the central dot to the time one of the RAs was entered. The maximum allowable RT was defined as 2000 ms; when subjects did not respond within this time the same trial was repeated. After each trial, an inter-trial-interval of 500 ms was presented. If during the acquisition phase the subject showed declining attention to the screen or was otherwise distracted the acquisition phase could be shortened (minimum amount of acquisition trials was set at 30).

Test phase

After acquisition, the test phase followed directly (32 trials) (see Figure 1). A test trial started with a similar dot with superimposed line drawing as in the acquisition phase, again serving as a start- and fixation-stimulus. However, after the subjects fixated on the dot (fixation time again varied randomly between 150-350 ms), the dot stayed on the screen for another 200 ms during which one of the effect-sounds that was previously triggered by one of the two eye-movements, was played after which the dot immediately disappeared. Then two identical 5.3° by 5.3° images of the same face (again randomly selected without replacement from Nottingham scans' emotional faces database) appeared 9.7° to the left and right of the center of the screen. The images were identical to avoid any influence on the subject's gaze preference. To further minimize influence on preference, the faces now pulsated in synchrony, they either both started growing or shrinking (randomized and with the same motion parameters as in the acquisition). Again, this was expected to evoke a spontaneous saccade and the question of interest was whether the direction of this saccade would be biased by the tone. Saccades towards the location that previously produced the tone were considered congruent while saccades towards the alternative location were considered incongruent. The minimum amount of test trials to enter analysis was 21. Except for absence of the effect after the saccade, the remaining procedure was as in the acquisition phase.

Awareness

After the experiment, adults were asked if they noticed any regularity in the sound mapping in the experiment. If so, they were asked what it was (e.g., "When I looked to the right I heard sound x, when looking to left I heard sound Y"). Then, all subjects were asked whether they noticed that there were two parts in the experiment. If they did notice, they were asked more specifically if they noticed any regularity in the sounds during the first (acquisition) phase, if not they were scored as unaware. If they noticed two phases but no regularity in the sound mapping, they were asked specifically if they had noticed that during the first phase there was a mapping between sounds and direction of looking. If they did they were considered “aware”, otherwise “unaware”.

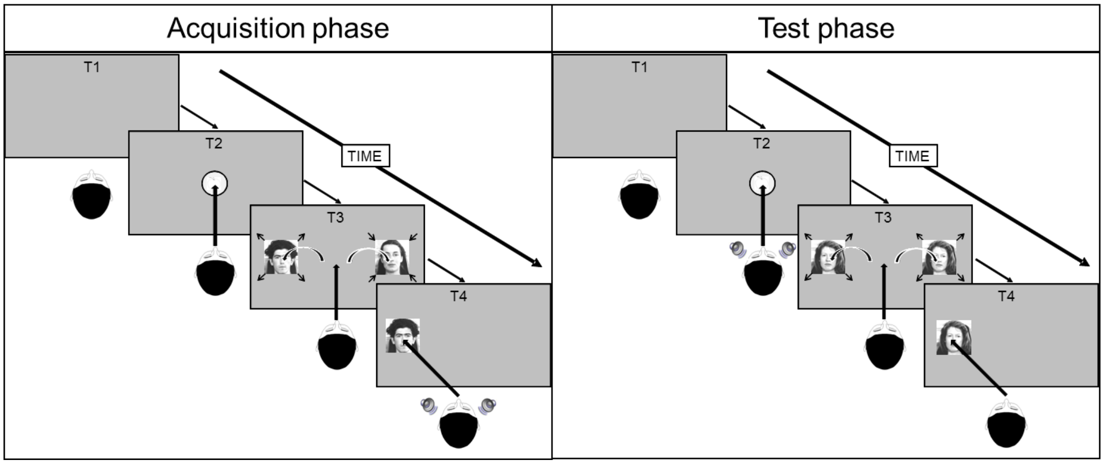

Figure 1. **Acquisition trial**. T1: Each trial starts with an intertrial interval of 500 ms. T2: A fixation dot is displayed at screen center. T3: After succesful fixation, faces appear at either side of the screen where they started to pulsate. T4: Depending on the saccade target, the face at the other side disappears and an effect sound is played for 200 ms. **Test trial**. T1: Each trial starts with an intertrial interval of 500 ms. T2: A fixation dot is displayed at screen center. After succesful fixation one of the previous action effects is played. T3: The dot diasppears whereafter the same face appears on both sides. T4: The participant freely chooses where to saccade.

Data acquisition

E-prime® was used to collect RTs during acquisition and test phases, the number of left and right responses during acquisition, and the number of congruent and incongruent responses during test. Furthermore, the E-gaze data files produced by E-prime® were imported into BrainVision Analyzer software (Version 1.05, BrainProducts, Germany) to analyze gaze position and pupillary data. First pupil sizes of both eyes were averaged to create more stable data. Artifacts and blinks as detected by the eye-tracker were corrected by using a linear interpolation algorithm. After this a 10 Hz low-pass filter was used, commonly used for pupil data (e.g., Hupe et al., 2009). To ensure there were no erroneous pupil data we then rejected artifacts using the parameters of a minimal pupil size of 1mm and a maximum of 5mm, furthermore the maximum allowed change of pupil-size was defined as .03mm in 17ms.

Given that the acquisition of action-effect associations is sensitive to the same factors as stimulus-response learning (Elsner & Hommel, 2004), the bias to respond in either direction during acquisition was calculated ("Acquisition Bias", AB=the number of leftward saccades minus the number of rightward saccades). As the size of this bias represents the degree to which participants were selectively exposed to one of the action effects, the AB variable was used as covariate in the analyses when appropriate.

Results

Acquisition phase

RT and response frequency

All ANOVA's were performed with age group as a between subjects factor. ANOVA's on the percentage of left responses and number of completed acquisition trials showed no effects, *p*'s > .3 (see Tabel 1).

Another ANOVA on mean RT revealed a significant effect of age group, $F(2,56) = 55.30$, $p < .001$, η2p = .66: Post-hoc Tukey HSD comparisons revealed that adults responded significantly faster than infants, *p*'s < .001) (see Tabel 1).

A repeated-measures ANOVA on RT's with right-vs-left as within subjects factor showed no effects, *p*'s > .8.

*Table 1***.** Mean scores of acquisition phase (standard deviation in brackets).

AGEGROUP SCORES	Number of acquisition trials	Percentage of left responses	RT in ms
7-month-olds	47.1 (3.36)	60.3 (37)	441 (49)
12-month-olds	47.15 (2.30)	48.9 (30)	440 (57)
Adults	48 (0)	52.4 (11)	293 (50)

Awareness

Only 4- of 24 adult participants reported being aware of the action-effect mapping in the acquisition phase.

Test phase

Response frequency

Again all ANOVA's were performed with age group as a between subjects factor. An ANOVA on the number of completed test trials showed a significant effect of age group, $F(2,56) = 7.71$, $p = .001$, η2p = .22. Post-hoc Tukey HSD comparisons revealed that adults and 7-months-olds completed more test trials than 12-months-olds (mean adults = 32, 7-months-olds = 31.6, 12-months-olds= 29.6, *p*'s < .02). This is probably due to increased agility and fussiness in the 12-months-olds and increased motivation in the adults (see Table 2).

A repeated measures ANOVA on percentage of responses with left vs. right as within subjects factor revealed that, overall, participants showed no tendency to saccade more often to either side, $p > .05$, however it did show an interaction with age group $F(2,52) = 4.11$, $p = .02$, η2p = .14. Separate comparison showed that while adults made more leftward saccades, $F(1,23) = 4.23$, $p = .05$, η2p = .16 (18 left vs. 14 right), 12-month-olds did the opposite $F(1,18) = 5.94$, $p = .03$, η2p = .25 (12 left vs. 18 right) (see Table 2).

More importantly for our purposes, a repeated-measures ANOVA on response frequency with congruency as within subjects factor showed that the percentage of acquisition-congruent vs. incongruent responses did not differ and congruency did not interact with age group, *p*'s > .6.

We additionally performed a median split on RT's on each subject, classifying trails as either fast or slow, and calculated the percentage of fast congruent responses vs. the percentage of slow congruent responses. We then performed a repeated-measures ANOVA on the percentage of congruent responses, with fast vs. slow as within subjects factor. We found no main effect of fast vs. slow on percentage (*F* > 1), we did find a significant interaction of fast vs. slow with age group *F*(1,56) = 3.58, *p* = .03, η2p = .11. We then tested the age groups separately, showing that adults had a higher percentage of congruent responses in their fast responses as compared to their slow responses (54 vs. 45%) *F*(1,23) = 11.5, *p* = .003, η2p = .33, while the infants showed no such effect (see Table 2).

We also performed a repeated-measures ANOVA on percentage of congruent reactions with Time (dividing the responses in three bins; trial1 - 10, 11 - 21 and 22-32) as within subjects factor which did not yield any effects (*p*’s > .2).

*Table 2***.** Mean frequency scores of test phase (standard deviation in brackets).

AGEGROUP SCORES	Completed test trials	Percentage of left responses	Percentage of congruent responses	Percentage of congruent responses in fast reactions	Percentage of congruent responses in slow reactions
7-month-olds	31.6 (1.55)	61.5 (34.9)	48.8 (6.9)	46.3 (12.4)	51.3 (12.6)
12-month- olds	29.6 (3.44)	38.1 (28.7)	49.8 (6.6)	51.8 (6.7)	47.6 (13.3)
Adults	32 (0)	56.5 (15.5)	49.7 (10.9)	54.4 (13.5)	45.1 (12.4)

Reaction times

Again all ANOVA's were performed with age group as a between subjects factor. Since the test-phase was self paced we also performed an ANOVA on inter-trial interval (ITI) and found a significant effect, $F(2,56) = 36.53$, $p < .001$, η2p = .57. Post-hoc Tukey HSD comparisons revealed that adults responded significantly faster than infants, *p*'s < .001 (see Table 3).

As in the acquisition phase, an ANOVA on RT's showed that adults responded faster than the two infant age groups, $F(2,56) = 89.07$, $p < .001$, η2p = .76; all HSD *p*s < .001.

Another repeated measures ANOVA on RT's with left vs. right as within subjects factor showed no effect, $p>.1$ (see Table 3).

More importantly for our purposes, we performed repeated measures ANOVA on RT's with congruent vs. incongruent as within subjects factor, using AB as a covariate which showed that acquisition-congruent responses were initiated 14 ms faster than incongruent responses, $F(1,55) = 4.20$, $p = .05$, η2p = .07, and this effect interacted with age group, $F(2,55) = 4.38$, $p = .02$, η2p = .14. Separate comparisons showed that the congruency effect was significant in adults, $F(1,22) = 10.60$, $p = .004$, η2p = .33, and 12-month-olds, $F(1,18) = 8.51$, $p = .009$, η2p = .32, but not in 7-month-olds, $F(1,13) = 2.51$, $p = .14$ (see Figure 2). Additional non parametric analysis in the 7-month-olds also failed to show an effect of congruency on RT's in this group (see Table 3).

We also performed a repeated measures ANOVA on RT's with Time (dividing the responses in three bins; trial1 - 10, 11 - 21 and 22-32) and congruence as within subject factors using AB as a covariate. We found an overall tendency regarding the main factor of Time ($F(2,102) = 2.64$, $p = .08$, η2p = .05) with slower responses as the test progressed which interacted with age group $F(4,102) = 3.72$, $p = .01$, η2p = .13, further separate testing revealed that only the 12-month-olds showed a significant slowing as the test progressed, *F*(2,32) = 7.02, *p* = .003, η2p = .31. No further interactions with Time were found *p*'s > .3. The main effect of congruency on RT's was significant *F*(1,51) = 10.85, *p* = .002, η2p=.18, and showed that congruent responses were initiated 23ms faster. The interaction of congruency with age group on RT's also reached significance *F*(2,51) = 5.70, *p* = .006, η2p = .18. Separate testing for the age groups revealed that the effect was significant in adults *F*(1,21) = 7.45, *p* = .013, η2p = .26, and in the 12-month-olds, *F*(1,16) = 13.90, *p* = .002, η2p = .47, but not in 7-month-olds, *F*(1,12) = .95, *p* = .35 (see Table 3).

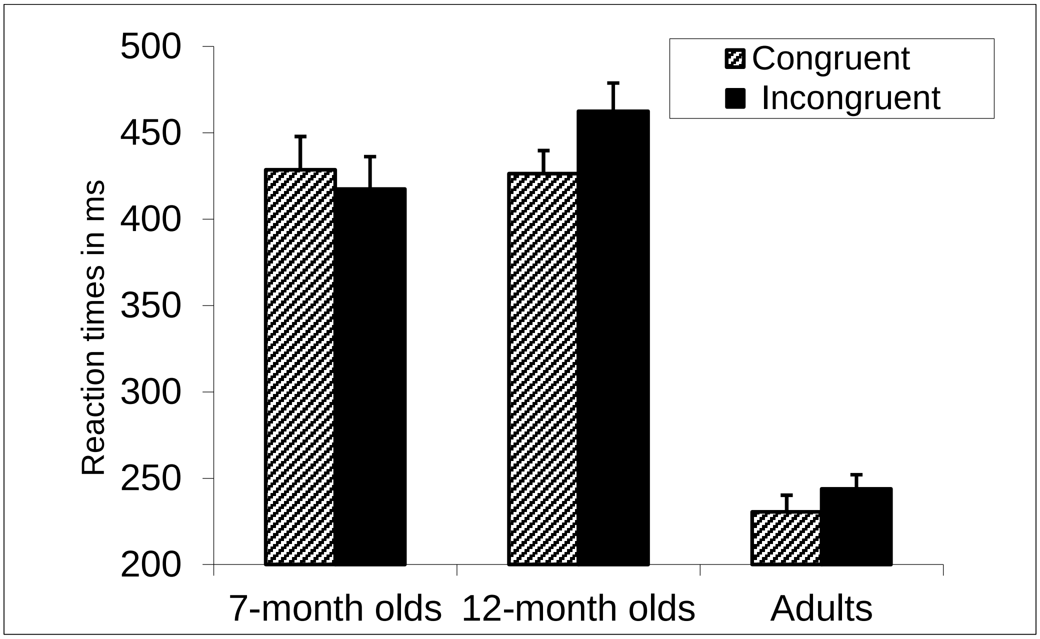

Figure 2. Mean reaction times (+SE) for adults (N=24) 7-month-olds (N=17) and 12-month-olds (N = 22) in congruent and incongruent test trials.

*Table 3***.** Mean RT scores of test phase (standard deviation in brackets).

AGEGROUP SCORES in ms	ITI	RT	RT Congruent	RT Incongruent	RT trial 1 -10	RT trial 11-21	RT trial 22-32
7-month-olds	2761 (842)	432 (71.9)	437.6 (77.9)	425.0 (78.0)	410.1 (71.1)	427.8 (74.7)	426.9 (109.5)
12-month-olds	2807 (902)	449 (62.3)	427.4 (60.5)	470.6 (83.4)	420.9 (85.7)	431.2 (53.8)	493.6 (104.7)
Adults	1218 (255)	237.6 (42.9)	230.4 (47.4)	244.0 (39.2)	245.9 (47.6)	234.8 (49.4)	229.2 (49.2)

Pupil dilation

TEPRs were sorted according to congruency of the response and the stimulus- and response-locked time functions were averaged. Segments were created, depending on the analysis, from 2000ms before the presentation of the sound or RT to 8000ms after while allowing for overlapping segments. Following the method used by Bradley et al. (2008), pupil-diameter measurement began after the initial pupil reflex caused by the fixation stimulus. Visual inspection showed the light reflex to end around 500ms after effect presentation (see Figure 3). To accommodate for the variable RTs across age groups and conditions, we considered both stimulus-locked and response-locked TEPRs. TEPRs were calculated as the percentage of dilation relative to the baseline to make the data more comparable across age groups.

First we analyzed whether the percentage of trials rejected due to erroneous data points differed across age groups. An ANOVA on the percentage of kept trials yielded a reliable main effect of age group, $F(2,56)=4.30$, $p = .02$, η2p = .13 (average percentage of kept trials: 7-month-olds 92%, 12-month-olds 92%, adults 99%). Post-hoc Tukey HSD comparisons showed that in adults significantly fewer segments were rejected than in the 12-months-olds (*p* = .03), and the same tendency was visible in the comparison of adults and 7-months-olds (*p* = .08), an unsurprising observation given the differences in attentional resources between infants and adults.

The stimulus-locked analysis of TEPRs in congruent and incongruent trials used a 500 ms pre-effect baseline (Beatty & Lucero-Wagoner, 2000). A repeated measures ANOVA pupil dilations with congruency as within subjects factor revealed no a priori effects of congruence on baselines (-500 to 0 ms), *p*'s > .10. TEPRs start from 200 to 300 ms after stimulus onset and peak around 1200 ms post-stimulus (Beatty & Lucero-Wagoner, 2000) in the range of 500ms to 2000ms (Beatty, 1982). We therefore calculated the mean TEPRs for congruent and incongruent responses as the mean percentage of change from baseline to 500-2000 ms post effect onset.

A repeated measures ANOVA revealed that, overall, participants showed larger relative dilations in incongruent trials, *F*(1,56) = 6.80, *p* = .01, η2p = .11, and this effect was not modulated by age group, *p* > .10 (see Figure 3). To have a closer look into developmental changes, we then analyzed the infant data separately. On average, infants showed larger relative dilations in incongruent trials, *F*(1,33) = 6.78, *p* = .02, η2p = .17, and this effect was not modulated by age group, *p*>.10. Of particular importance (given the reaction time results), the congruency effect remained significant when the 7-months-olds were tested separately, *F*(1,14) = 12.0, *p* = .004, η2p = .46 .

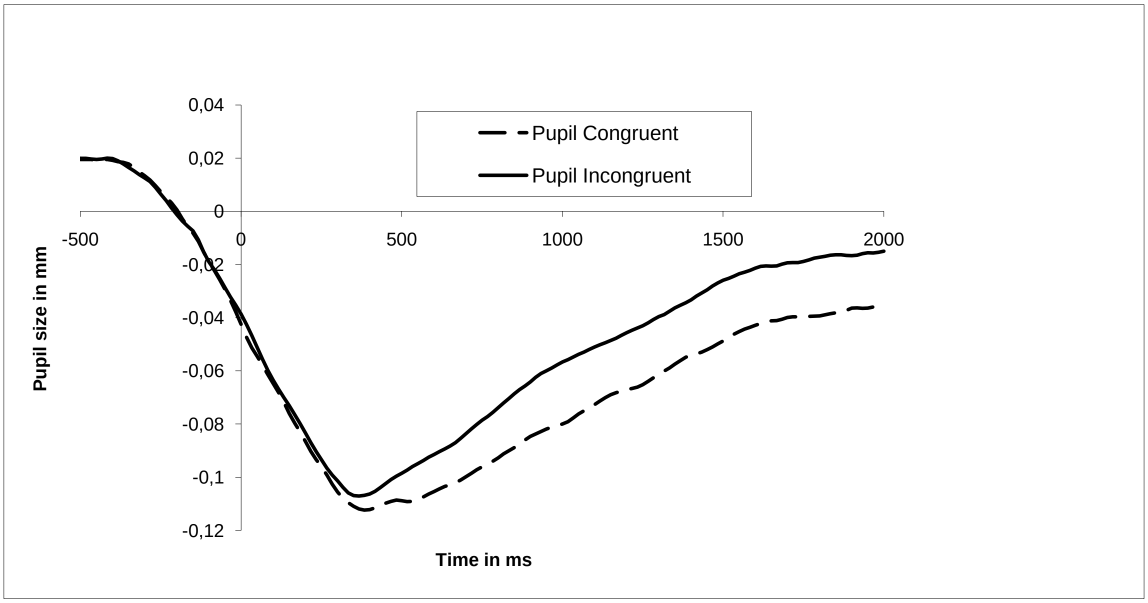

Figure 3. Mean relative pupil sizes for congruent and incongruent responses, stimulus-locked.

For the response-locked analysis, we calculated the percentage of dilation from a 700-ms time window from saccade onset on, to a 200-ms pre-response baseline. A repeated measures ANOVA showed no a priori effects of congruence on baselines (-200 to 0 ms), *p'*s > .10. The analysis of these data yielded a significantly larger relative dilation in incongruent than congruent trials, *F*(1,56) = 7.82, *p* =.007, η2p = .12, while the interaction with age group was not significant, *p* > .10 (see Figure 4). A separate analysis of the infant data showed a main effect for congruency, *F*(1,33) = 8.41, *p* =.007, η2p = .20, that was not modulated by age group, *p*>.10.

Another version of this analysis with a 1000-ms pre-response baseline produced a different pattern (a repeated measures ANOVA revealed no a priori effects of congruence on baselines *p'*s > .10): a congruency effect, *F*(1,56) = 10.19, *p* =.001, η2p =.17 (see Figure 4), but also an interaction of congruency with age group, *F*(2,56) = 3.99, *p* =.02, η2p =.13 . Separate testing showed that the effect was only reliable in the 7-months-olds *F*(1, 14) = 10.59, *p* =.006, η2p = .43, while the other two groups did not reach significance.

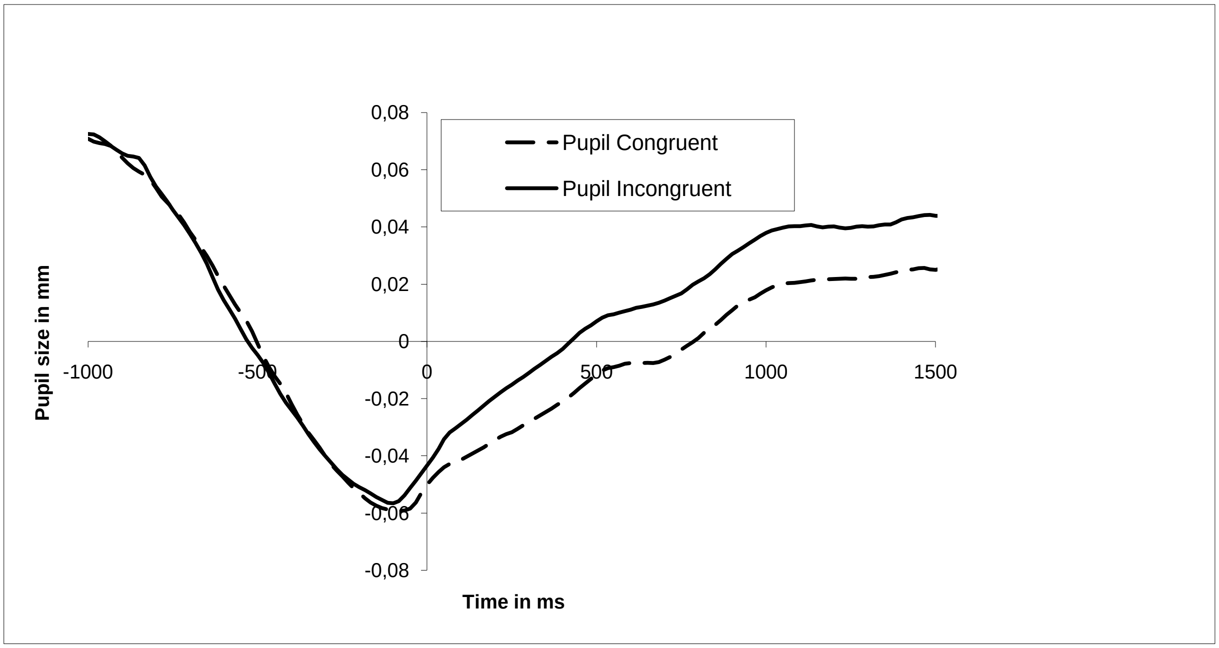

Figure 4. Mean relative pupil sizes for congruent and incongruent responses, response-locked.

Chapter 7

Discussion

The aim of the current study was to directly compare action-effect learning in infants and adults using a novel paradigm that relies on oculomotor actions that occur spontaneously and do not require verbal instruction. We succeeded in developing an eye-tracking paradigm that was equally suitable for both very young infants and adults. Moreover, the paradigm allowed for concurrently investigating the impact of action-effect learning on biases in, and the efficiency of action selection, as measured by response choice and RT, respectively, and on action effort and/or monitoring, as indicated by pupil dilation.

As expected from ideomotor theory (James, 1890; Hommel et al., 2001), adults and 12-months-olds were faster to carry out responses that were congruent with the present trigger tone (i.e., responses that produced this tone in the acquisition phase) than incongruent responses. The only difference between congruent and incongruent responses was their past relationship with the tones, which indicates that the congruency effect reflects associative knowledge acquired during the acquisition phase. Moreover, the fact that the tones now primed the response they previously had followed suggests that the underlying association was bidirectional in nature. Both observations are consistent with ideomotor theory and fit well with the observations of Herwig & Horstmann (2010), who reported oculomotor action-effect learning in adults. Interestingly, these authors used visual action effects while the present study employed auditory effects. This confirms that the mechanism underlying action-effect learning is general and not bound to a particular modality, as long as the effects are contingent on, and temporally close to the corresponding actions (Elsner & Hommel, 2004). We also found that most adults were unaware of the saccade-effect mapping, which is in line with the idea that action-effect acquisition is a low-level, fast and automatic process that does not require attention. It seems reasonable to assume that the same holds for the infants.

The present findings fit with observations from manual actions in a developmental study of action-effect acquisition by Verschoor et al. (2010). In this study, reliable RT effects were found in 9-, 12- and 18-months-olds, indicating action-effect acquisition in these age groups. We obtained a similar RT effect in 12-month- but not 7-months-olds. We take this to imply that, although ongoing contingent action-effects can influence behavior and memory at this age (Gergely & Watson, 1999; Rovee-Collier, 1999), action-effect associations cannot yet be reversed to play an active role in prospective action control. This is in line with the dissociation between acquired action knowledge and the use of such knowledge found by Keen (2005) and Sommerville, Woodward, and Needham (2005). It also fits with similar dissociations in looking-time studies wherein infants are reported to possess knowledge on actions they cannot perform themselves (e.g. Verschoor & Biro, 2012; Csibra et al., 1999).

Although we found an effect in the adults indicating that fast responses were more likely to be acquisition-congruent than slow responses (this can be taken as further evidence that action-effect learning relies on a fast and automatic mechanism, at least in adults), congruency effects were restricted to RT’s and did not affect response choice. One might assume that the lack of frequency effects suggests different developmental pathways with respect to manual and oculomotor actions. There are several arguments against this interpretation. For one, although manual free-choice studies have shown that presenting an action effect can bias response choice towards the response that had previously produced that effect in adults (e.g., Elsner & Hommel, 2001; Eenshuistra et al., 2004; Kray et al., 2006), even in free-choice studies that did find a reliable effect on response frequency, frequency turned out to be less sensitive to action-effect learning than reaction time. For instance, Verschoor et al. (2010) obtained a congruency effect on response frequency in 18-month olds, but not in younger infants, while congruency affected reaction time in 12- and 9-month-olds as well. Since Verschoor et al. (2010) used only very few test trials, one might suggest that in their study extinction, which younger infants are more susceptible to (e.g. Hartshorn et al., 1998), could not have played a major role. In the present study the test phase contained considerably more test trials (which were necessary to get sufficiently clean pupil-dilation data). Our paradigm thus provided more opportunity for extinction since action effects were no longer presented during test-trials. However, we tested whether the effect of congruency on RT’s and response frequency declined over time and found no such effect. Even though action-effect learning can be demonstrated under extinction conditions in principle, extinction does make the effect weaker (Elsner & Hommel, 2001)—and it may have weakened it enough to selectively annihilate the frequency effect altogether. Moreover, Herwig and Horstmann (2011) showed under extinction conditions a reliable reaction-time effect in their very similar, albeit forced-choice occulomotor paradigm using even more test trials (32 vs. 96). This indicates long-lasting, extinction-resistant bidirectional associations.

For another, since the current paradigm is conceptually very close to that of Herwig and Horstmann (2011), and the adults failed to show frequency effects, it is more likely that in the manual version of the action-effect task by Verschoor et al. (2010), the action effects affected response choice differently from the current paradigm. In manual action-effect paradigms the only attention-drawing events in the test phase are the presented action effects. Its mere presence is unlikely to affect action choice directly, so that all possible response biases can be attributed to the degree to which the action effect reactivated a previously acquired association, which then spread activation to the corresponding response representation. In other words, even though action effects attract exogenous attention, they eventually impact action selection in an entirely endogenous fashion. Indeed, neuroimaging studies have shown that the presentation of previously acquired action effects activates the supplementary motor area, which underlies endogenously-driven but not exogenously-driven action selection (Elsner et al., 2002; Melcher, Weidema, Eenshuistra, Hommel & Gruber, 2008; Paulus et al., 2012; Paulus, Hunnius & Beckering, 2013). In contrast, in our oculomotor version of the task, the endogenous impact of the action effect competes with the direct, exogenous impact of the saccade goals—the faces in our case. It is possible that this exogenous impact is so strong that it outweighs the impact of the endogenous bias to a degree that the latter is too weak to determine which response is being chosen, even though it can still speed up congruent and/or slow down incongruent responses. Accordingly, the present findings do not necessarily require the assumption that action-effect learning is different in, or follows different developmental pathways with respect to manual and oculomotor actions. One could even speculate that the exogenous attention evoked had a stronger impact on the 7-months-olds, thus resulting in a lack of RT effect in this group.

As expected, we found reliable effects of congruency on pupil dilation with incongruent saccades resulting in larger relative dilations. These findings need to be interpreted with caution, as there are no shared standards regarding the handling of TEPRs. TEPRs can start from 200-300ms after stimulus onset and peak around 1200 ms post-stimulus (Beatty & Lucero-Wagoner, 2000), sometimes even later (e.g., Bradly et al., 2008; Beatty, 1982). The effects found do fit within these temporal dynamics. On the other hand, little is known about developmental and aging-related changes in these biometric variables across the lifespan. Since our experiment is self-paced (to ensure infant cooperation) ITI’s do vary systematically between age groups, with shorter ITI’s for the adults. What’s more, in the 12-months-olds and adults RT’s also vary with congruency. These timing factors could reduce pupillary effects in these age groups. Alas, due to the attentional abilities of infants standardizing the ITI’s was not an option. Furthermore, since light adaptation dilations decrease in amplitude and latency with increasing gestational age (Cocker, Fielder, Moseley, & Edwards, 2005) one could speculate that the same could hold true for TEPR’s. Thus ideal intervals for measuring TEPR’s and baselines may vary accordingly. Therefore, pupillary effects should be expected to be most pronounced in the 7-months-olds. This is indeed what we find. In the current study we chose intervals as suggested by the literature. However, in the literature there is no standard for response related evaluative effects.

To accommodate for the variable RTs and ITI's in the current experiment, we considered both stimulus-locked and response-locked TEPRs, which however yielded identical outcomes. Of particular interest, both analyses revealed main effects of congruency but no interaction with age. Moreover, the congruency effects remained reliable when being tested in the 7-months-olds alone, the only age group that did not show a congruency effect in RTs. On the one hand, the fact that 7-months-olds are sensitive to the congruency between their action and the presented action effect demonstrates that they have acquired information about the relationship between their actions and the novel auditory effects these actions produced in the experiment. Accordingly, we take this observation to indicate that even the youngest group was able to integrate some kind of information about actions and their effects. On the other hand, the dissociation between the dilation effect and the RT effect in these infants suggests that the two measures do not assess the same underlying processes.

As suggested by Band et al. (2009) and Blakemore, Frith and Wolpert, (1999), action-effect associations may not only serve as an informational basis for action selection, the major theme of ideomotor theory, but also for predicting the perceptual consequences of an action. This prediction can be matched against the actually produced consequences in order to evaluate whether the action goal was reached. The late timing of our stimulus- and response-locked congruency effects in the TEPRs suggests that these effects were picking up processes related to action evaluation rather than action selection proper. Indeed, error-related-negativity-type patterns as reported by Band et al. (2009) in connection to action prediction have been found to be related to pupillary responses as well (Wessel, Danielmeier & Ullsperger, 2011). These pupillary effects might thus indicate a mismatch between expected and actual action effect in incongruent trials and/or reflect the thereby triggered adaptive processes that update the system's knowledge about action-effect relationships. The first possibility would fit well with the violation-of-expectation approach suggested by Jackson and Sirois (2009), and Gredebäck and Melinder (2010), while the interpretation in terms of mismatch-induced control processes would be more along the lines of the traditional TEPR literature—which focuses on arousal, attention allocation, cognitive load and mental effort. In any case, the effect reflects knowledge about action-effect contingencies, and our stimulus- and response-locked findings suggest that this knowledge is equally present in all three age groups.

This dissociation between RT findings, which imply action-selection effects in adults and 12-months-olds, and TEPRs, which suggest action-evaluation effects in all participants, allow for two important conclusions. First, the processing of an action-effect stimulus activates a representation that creates particular expectations, without necessarily activating the corresponding actions. This means that action-effect expectation may be correlated with, and perhaps even functional for action selection (Kühn, Keizer, Rombouts, & Hommel, 2011) in older agents, but raising an expectation is not identical to selecting an action. Second, infants' abilities to construct action-effect expectations develop earlier than their abilities to use action-effect representation for intentional action selection. A similar dissociation between acquisition of action-effect knowledge versus use of action-effect knowledge was reported by Sommerville et al. (2005). They showed that violation-of-expectation to a change of goal was influenced by action-effect experience in 3-months olds, while the observation of actions and their effects did not influence action production.

There are several possible reasons for why selection abilities develop more slowly and why action-effect knowledge in the 7-months-olds affected expectation-related effects only. One possibility is that associations between motor patterns and novel action effects are either not yet bidirectional, too weak, or take too much time to retrieve to affect performance under our testing conditions. Another possibility is that novel action effects are not yet directly associated with actions but only with representations of already existing action effects. Taken altogether, it seems safe to assume that, at 7 months of age, knowledge about relations between actions and their effects has a stronger impact on the prediction of action effects than on the selection of intentional actions.

To conclude, the dissociation we obtained in 7-months-olds suggests a developmental precedence of action monitoring over intentional action selection. This again suggest that infants acquire the knowledge necessary for performing intentional actions sometime before they have (fully) developed the cognitive machinery necessary to make use of that knowledge to perform intentional action (Keen, 2005; Sommerville et al., 2005). Combining the current data with those of Verschoor at al. (2010) suggests a major change in action-effect learning from just action monitoring to action selection, just before the ninth month of age. If one takes into account the functional and representational equivalence of self-performed and perceived actions as suggested by Theory of Event Coding (Hommel et al., 2001), this pattern fits with data suggesting that at 6 months of age infants can understand goal directed action (e.g. Woodward, 1998), or more accurately, experience violation-of-expectation to a change of goal, but are unable to perform true intentional action (distinguishing means from ends) until around 8 to 9 months of age (Goubet et al., 2006; Hauf, 2007; Piaget, 1936). Our findings also fit with results from studies on action perception, showing that infants at 9- but not 7-months of age can use observed action-effect relations to guide behavior (Hauf & Aschersleben, 2008). Additionally, our data suggest that motor resonance when listening to previously self-produced sounds in 8 months olds, as found by Paulus et al. (2012), might indeed reflect the existence of knowledge about action-effect relations; and yet, we do not necessarily expect this knowledge to result in overt behavior, at least not at 7-months of age. Similar evidence for action-knowledge activation during action observation has been obtained in infants as young as 6 months (Nyström, 2008). Some authors have argued that it is lacking representational equivalence between self-produced actions and observed actions that prohibits infants younger than nine months from imitation (Hauf, 2007). Our data, together with those of Paulus et al. (2012-2013), Verschoor et al. (2010), Nyström (2008), and Sommerville et al. (2005), suggest that it is not representational equivalence that is reached by 9 months of age, but the ability to successfully use bidirectional action-effect associations, learned either by observation or experience, for voluntary action.

Chapter 8: The developing cognitive substrate of sequential action control in 9- to 12-month-olds: Evidence for concurrent activation models

Verschoor, S. A., Paulus, M., Spapé, M., Biro, S. & Hommel, B

Abstract

Nine-month-olds start to perform sequential actions. Yet, it remains largely unknown how they acquire and control such actions. We studied infants' sequential-action control by employing a novel gaze-contingent eye tracking paradigm. Infants experienced occulo-motor action sequences comprising two elementary actions. To contrast chaining, concurrent and integrated models of sequential-action control, we then selectively activated secondary actions to assess interactions with the primary actions. Behavioral and pupillometric results suggest 12-month-olds acquire sequential action without elaborate strategy through exploration. Furthermore, the inhibitory mechanisms ensuring ordered performance develop between 9 and 12 months of age, and are best captured by concurrent models.

Chapter 8

Introduction

Infants are active, goal-directed agents (e.g., McCarty, Clifton, Ashmead, Lee, & Goubet, 2001). Interestingly, some of the actions they produce can be considered sequential, such as reaching for a rattle in order to shake it — a rather simple sequence, that comprises two dissociable components that differ in function and motor demands. Piaget (1936) and others (Claxton, Keen, & McCarty, 2003; Hauf, 2007; Willatts, 1999; Woodward and Sommerville, 2000; Woodward et al., 2009) have stated that true goal-directed action emerges around 9 months of age when infants begin to be able to organize means-end action sequences in the service of overarching goals. Yet, the cognitive substrate of early sequential action control in infants remains completely uncharted territory. The purpose of the current study is to explore the cognitive mechanism sub-serving sequential action control in infants.

Development of action control in infancy

There are three prerequisites for infants to control sequential action: that they can represent actions, that they can represent sequential information, and that they can combine those abilities to represent and control sequential action. Let us turn to the first prerequisite. There is ample evidence that actions are represented in terms of their effects. In his ideomotor theory, James (1890) states that actions are learned on the fly through sensorimotor exploration; an automatic mechanism creates bidirectional associations between perceived effects and the actions producing them (Hommel, 2009; Hommel et al., 2001). These associations bring the actions under voluntary control, enabling the agent to activate the action by “thinking of” the corresponding effect. The theory can thus account for learning new actions and new goals.

This idea is typically tested in a two-stage paradigm. Experimenters first let subjects perform actions that lead to specific effects. After acquisition, they test if exogenously cueing an effect cues the action that previously caused it (Elsner & Hommel, 2001; Greenwald, 1970). This approach resulted in demonstrations of bidirectional action-effect acquisition for a wide range of actions and effects in children (Eenshuistra et al., 2004; Kray et al., 2006) and adults, suggesting the mechanism responsible to be fast-acting (Dutzi & Hommel, 2009), automatic (Elsner & Hommel, 2001; Band et al., 2009), implicit (Elsner & Hommel, 2001; Verschoor, Spapé, Biro & Hommel, 2013), and modulated by the same factors that influence instrumental learning (Elsner & Hommel, 2004) (for a review on action-effect learning see: Hommel & Elsner, 2009). Furthermore, action-effects have also been found to be important for action evaluation (Band et al., 2009; Verschoor et al., 2013).

Infant research has mainly been restricted to the importance of action effects for third-person action interpretation (e.g., Biro & Leslie, 2007; Hauf, 2007; Kiraly et al., 2003; Paulus, 2012; Paulus, Hunnius, & Bekkering, 2013; Woodward, 1998, for a review, see: Hauf, 2007; Kiraly et al., 2003) and imitation (Hauf & Aschersleben, 2008; Klein et al., 2006; for a review see: Elsner 2007; Paulus, 2014). Such findings are corroborative in view of the upsurge of theories stressing similar representations for first- and third-person action (e.g. Baker, Saxe & Tenenbaum, 2009; Fabbri-Destro & Rizzolatti, 2008; Melzoff, 2006, 2007; Tomasello, 1999). Interestingly, increased model- to self- similarity aids imitation (Shimpi, Akhtar & Moore, 2013). Yet given their focus on action understanding, such studies tell us little about the function action effects have for the development of action control in infancy.

Direct evidence regarding action-effect learning was recently obtained from first-person paradigms similar to that of Elsner and Hommel (2001). Verschoor et al. (2013) showed that 7-month-olds use action effects for first-person action monitoring. By eight months, infants show motor resonance when listening to previously self-produced action-related sounds (Paulus, Hinnius, Elk & Beckering, 2012). The youngest infants showing evidence for reversing bidirectional action effects for action control are 9-month-olds (Verschoor et al., 2010). Comparable results were found in 12- (Verschoor et al., 2013), and 18-month-olds (Verschoor et al., 2010). Additionally 6-, 8- (Wang et al., 2012) and 10-month-olds (Kenward, 2010) anticipate action outcomes. Taken together these studies illustrate that 7-month-olds represent and monitor first- and third-person action in terms of action effects, while 9-month-olds additionally use action effects for action control.

Representing sequential information in infancy

Another prerequisite for representing sequential action is the ability to encode sequential information. Infants can register whether items are consistent with familiarized deterministic or probabilistic sequences (Romberg & Saffran, 2013). For instance, infants are susceptible to sequential grammar information in speech from birth (Gervain, Berent, & Werker, 2012; Teinonen, Fellmann, Näätänen, Alku & Huotilainen, 2009), 3-month-olds are susceptible to spatiotemporal (Wentworth Hait & Hood, 2002) and audio-visual sequences (Lewkowicz, 2008) and 8-month-olds to analogous information in artificial sound (Marcus, Fernandes & Johnson, 2007). Studies like these suggest an implicit, early-appearing, domain-general statistical information-acquisition mechanism for sequential information (e.g. Kim, Seitz, Feenstra & Shams, 2009; Kirkham, Slemmer & Johnson, 2002; Marcovitch & Lewkowicz, 2009) thought to sub-serve action- and language-segmentation (e.g. Baldwin, Andersson, Saffran & Meyer, 2008; Saffran, Johnson, Aslin & Newport, 1999). Nonetheless these studies leave open whether infants encode ordinal information among sequence elements. Indeed, Violation Of Expectation (VOE) research suggests that while 4-month-old infants encode statistical sequential properties, they cannot code the invariant order of sequences (Lewkowicz & Berent, 2009). This ability emerges during the second half of the first year (Brannon, 2002; Picozzi, de Hevia, Girelli & Macchi-Cassia, 2010; Suanda, Tompson & Brannon, 2008).

The reviewed literature shows that the first two prerequisites for infants' representation of sequential action emerge around 9 months. Yet, the question remains whether they can actually combine these abilities to represent and control action sequences. Indirect evidence comes from research that suggests infants are able to interpret third-person sequential actions. Evaluating such actions requires them to be parsed in order to perceive overall syntax and ultimately their goal (Conway & Christiansen, 2001; Lewkowicz, 2004; Baldwin, Baird, Saylor, & Clark 2001). VOE studies report that around the age of 6 months infants start to evaluate the efficiency of sequential actions (Biro et al., 2011; Csibra, 2008; Gergely & Csibra, 2003; Verschoor & Biro, 2012) and causality towards their goals (Baillargeon, Graber, DeVos & Black, 1990; Woodward & Sommerville, 2000). Olofson and Baldwin (2011) found that 10-month-olds take into account the kinematics of an observed reaching motion to judge whether it is part of a familiar action sequence. Yet, Paulus, Hunnius, and Bekkering (2011b) showed that 20-, but not 14-month-old infants use such information to predict goals. Additionally, Gredebäck, Stasiewicz, Falck-Ytter, Rosander and von Hofsten, (2009) showed that 14- but not 10-month-olds' predictive eye movements are influenced by the models later intention with the object. Moreover, infants use social context to bind actions of two collaborating actors into action sequences for goal evaluation (Henderson & Woodward, 2011; Henderson, Wang, Matz & Woodward, 2013) and goal prediction (Fawcett & Gredebäck, 2013). Although these studies provide evidence that infants have some understanding of others' sequential action, they do not reveal the cognitive mechanisms underlying infants' control of their own sequential action

Turning to infants' own action control, studies on (deferred) imitation of *enabling* action sequences (sequences in which one action is temporally prior to and necessary for a subsequent action) report that only a subset of 9-month-olds can (immediately) reproduce such sequences under ideal circumstances (e.g., Bauer, Wiebe, Waters & Bangston, 2001; Carver & Bauer, 1999, 2001). Addition of salient action-effects to separate action steps increases performance (Elsner, Hauf, & Aschersleben, 2007). However, production of sequential action is in itself not enough to evince infants' sequential action control, since subsequent actions may simply be subsequent. In enabling sequences, stimulus enhancement could externally trigger such sequences. Indeed, an advantage for imitating enabling- over arbitrarily-ordered actions is reported (e.g. Barr & Hayne, 1996; Bauer, Hertsgaard, & Wewerka , 1995; Mandler & McDonough, 1995, for a review see: van den Broek, 1997). Earliest evidence for imitation of arbitrarily-ordered action sequences is reported for 16-month-olds (Bauer, Hertsgaard, Dropik & Daly, 1998).

Advance planning would make a stronger point for sequential action control. Claxton et al. (2003) reported that 10-month-olds plan the kinematics of reaching depending on subsequent intentions. Furthermore, McCarty, Clifton and Collard (1999) showed that 19- but not 14-month-olds inhibit reaching for an object with their dominant hand when this is inefficient towards an overarching goal (see Cox & Smitsman (2006) for a conceptually similar result in 3-year-olds). Both McCarty et al.'s (1999) and Cox and Smitsman's (2006) tasks depend on inhibition of pre-potent responses and suggest inhibition is important for sequential action planning (for a review see McCormack & Atance, 2011). Likewise, the disadvantage for reproducing arbitrarily-ordered action sequences seems to come from an increased need to temporally organize such sequences (Bauer, Hertsgaard, Dropik & Daly, 1998), which many theorists hypothesize inhibition to be crucial for (e.g. Constantinidis, Williams & Goldman-Rakic, 2002; Norman & Shallice, 1986).

To sum up, the studies mentioned above suggest a rudimentary ability to control first- person sequential action emerges by the end of the first year. Nonetheless, these studies shed little light on the cognitive format of the representations themselves on which sequential action control operates. Understanding the cognitive ontology of such representations is essential in interpreting the results of studies in which sequential action is the subject matter.

Models of sequential action representation

As there is little specific developmental literature on the subject, we turn to general psychological theories on sequential action control. Through the years many influential theoretical incarnations of sequential action representation have been conceived (de Kleijn, Kachergis, & Hommel, 2014). All of these theories hold that sequential actions consist of elementary actions that are somehow combined into sequences, as suggested the observation that the speed of sequence-initiation increases with the number of elements therein (e.g., Henry & Rogers, 1960; Rosenbaum, 1987). The theories can be distinguished into three ontological types that differ with respect to the representations action control operates on. We refer to them as chaining, concurrent, and integrative theories of sequential action control (see Figure 1). Chaining theories stress that elementary actions are selected and combined through association processes. Concurrent (Hebbian) theories focus on competitive processes that account for the orderly production of an action sequence. Integrated approaches highlight crosstalk between elementary actions resulting in chunked actions.

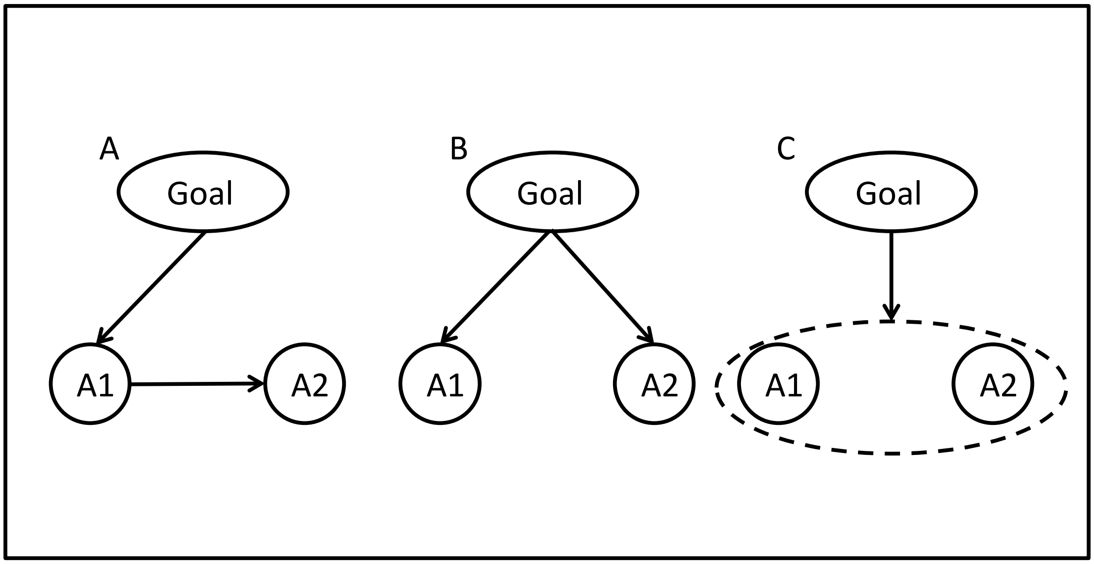

Figure 1. **Models of sequential action.** Schematic representation of activation in **A**: Chaining models of sequential action, activation cascades forward through the different elementary actions, **B**: Concurrent models of sequential action, all elementary actions are activated simultaniously whereafter competion through inhibition ensures the correct order of execution, **C**: Integrated models of sequential action, the sequence of actions has been integrated into a new elementary action

.

The prototypical theory of sequential action is James' (1890) chaining theory. It holds that elementary actions can be chained by sequentially activating the anticipated effect of each element. With practice, the sensory effect of each elementary action becomes associated with the next elementary action through stimulus-response learning, thereby eliminating the need for sequential activation. The model thus effectively reduces sequential action representation to a combination of ideomotor and stimulus-response learning. Furthermore, James' theory can account for the finding that infants better encode enabling- than arbitrarily ordered action sequences (Barr & Hayne, 1996; Bauer, Hertsgaard, & Wewerka, 1995; Mandler & McDonough, 1995), since the proposed feedback dependent effect-response learning he proposes is aided by stimulus enhancement in such sequences. Although James' theory is temptingly simple, it has a number of important drawbacks resulting in additions to the model. Hull (1931) pointed out that to stay goal-directed and flexible during performance, the representation of the end state should remain active during execution to compare the actual to the expected outcome. Hull thus introduces hierarchy into the representation by proposing continuous activation of an overarching goal. Secondly, in the conception of James (1890) the second action of a sequence is cued by the sensory effect of the first, suggesting sequential action to rely on sensory feedback. Yet, empirical evidence suggests feedback mechanisms to be too slow to account for the speed of practiced sequential action (e.g. Sternberg, Monsell, Knoll & Wright, 1978). Greenwald (1970) suggested that instead of the sensory effects, the anticipated action effects of the preceding action are associated to those of the next action. This enables the initiation of the sequence by anticipating its end effect. However the model does not specify how the end effect activates the sequence instead of just the final elementary action.

An important criticism on chaining models is that they seem to imply that elementary actions are equally associated with preceding and subsequent actions, making orderly performance of sequences impossible. In other words, chaining models of sequential action assume, but fail to describe how activation moves forward through the sequence. This lack of temporal dynamics in chaining models resulted in the emergence of a second ontological class of theories, the concurrent activation theories. Estes (1972) suggested an initial concurrent activation of all elementary actions by a superseding goal. Thereafter, temporal inhibitory processes ensure that activation moves forward through the sequence. To guarantee such forward flow he introduced inhibitory links flowing from each element to the next and secondary self-inhibition for completed elements (e.g., Henson, 1998, but also James, 1890), equivalent to inhibition of return in visual attention (Posner & Cohen, 1984; Houghton & Tipper, 1996). The concurrent model can account for the empirical finding of more prospective than retrospective intrusion errors (Dell, Burger, and Svec, 1997; Lindenberger, Rünger & Frensch, 2000; Rumelhart & Norman, 1982, for a review see: Houghton & Tipper, 1996). Concurrent models of sequential action representation thus utilize inhibition processes which are implicated in studies on action planning in infants (McCarty et al., 1999; Cox & Smitsman's, 2006).

The third class of theories, integrative theories of sequential action control, does not presuppose that action elements remain independent when combined into sequence. Such theories state that through practice elementary actions can be integrated into one common action plan or "chunk", implying considerable crosstalk between the elementary actions (Miller, 1956; Sakai, Kitaguchi & Hikosaka, 2003). Their main support comes from studies that find reductions in the sequence-length effect by extensive practice (e.g. Klapp 1995). Chunking of the elementary actions is possible by relating the sequences to internal or external context thus creating a unique identifying criterion for the associations (Hull, 1931). This Integration process can account for crosstalk between elementary actions and thus explains end-state comfort effects (Rosenbaum et al., 1990) as found in infant studies (Claxton et al., 2003; Cox & Smitsman, 2006; McCarty et al., 1999).

Experimental approach

Chaining, concurrent and integrated models generate different predictions with regard to the spreading of activation from one sequence element to another. Consider a sequence of two elements, with element A preceding element B. All models imply that priming or otherwise activating A might spread activation to B, but they differ regarding their predictions when B is primed/activated. James' (1890) chaining model would not predict that priming B leads to activation of A, since the sequence is assumed to be represented by unidirectional effect-response bindings (A→B). Greenwald's (1970) version would predict the spreading of activation from B to A, as sequences are represented by associations formed between the effects of the elementary actions. Conversely, concurrent activation models would predict that activating B leads to the inhibition of A, as activation is allowed to spread in forward direction only and backward connections are inhibitory. Finally, integrated models would predict that activating one element would activate the representation of the entire sequence, including A. The aim of the present study was to pit these different predictions against each other.

In the current study we were not only interested in the cognitive substrate of sequential action control, but also in the development thereof. Given findings that infants develop the ability for sequential action control around the end of the first year of life (e.g., Claxton et al., 2003), we hypothesized to find a developmental change in the cognitive substrate of sequential action control between 9 and 12 months of age. This line of reasoning is also supported by findings that the prerequisites for the ability seem to emerge in this interval. In 9-month-olds the ability to represent sequential information and action is rudimentary at best. Nonetheless, and crucial to the experimental logic, 9-month-olds represent actions in terms of their effects.

To tackle our questions regarding sequential-action control, we modified a recently developed gaze-contingent eye-tracking paradigm that assessed action-effect learning in infants and adults (Verschoor et al., 2013). This paradigm, conceptually identical to that of Elsner and Hommel (2001), overcomes problems arising due to limited motor control in infants (Verschoor et al., 2013; Wang et al., 2012). Verschoor and colleagues (2013) first let subjects perform actions that lead to specific effects. After acquisition, they tested whether exogenously cueing the effects primes the action that previously caused it. The paradigm uses eye movements which infants can accurately control from 4 months of age (Scerif et al., 2005) and which can be considered goal-directed (Gredebäck & Melinder, 2010; Falck-Ytter et al., 2006; Senju & Csibra, 2008). The paradigm records Reaction Times (RTs) and Response Frequencies (RFs). The study of Verschoor and collegues (2013), and other recent studies that demonstrated saccade-effect learning in adults (Huestegge & Kreutzfeldt, 2012; Herwig & Horstmann, 2011), showed shorter RTs for responses congruent with the previously acquired action-effect association (Verschoor et al., 2013). There is strong evidence that RT and RF differ in their sensitivity to congruency effects depending on age. In 9- and 12-month-olds RT is a sensitive measure (Verschoor et al., 2010; Verschoor et al., 2013), while in 18-month-olds RF additionally diagnoses congruency effects (Verschoor et al., 2010). The paradigm concurrently records Task-Evoked Pupillary Responses (TEPRs). The use of TERPs is relatively new in developmental research (Falck-Ytter, 2008; Jackson & Sirois, 2009; Laeng et al., 2012; Verschoor et al., 2013). TERPs indicate motivational phenomena such as increased arousal (Bradley et al., 2008; Laeng & Falkenberg, 2007), attention allocation (e.g. Hess & Polt, 1960), cognitive load (Kahneman & Beatty, 1966) and mental effort (Kahneman, 1973; Hess & Polt, 1964). Whatever the exact interpretation of the measure, it enables us to contrast acquisition contingent vs. non-contingent responses since all interpretations suggested that dilations should be larger for actions requiring more processing. Furthermore, pupil TERPs are sensitive to congruency in all ages, showing lesser dilation during congruent action (Verschoor et al., 2013). Thus, given that we tested 9- and 12-month-olds, we mainly expected congruency effects on RTs and TERPs.

In previous studies (that all used single-component actions), the definition of congruency was straightforward: participants were exposed to two action-effect contingencies during acquisition, in which responses A and B were followed by action effects 1 and 2 (A□1; B□2). Performing action A in response to (or as a result of being primed by) action effect 1 in the test phase (1□A) would be considered congruent, while performing the same action in response to effect 2 (2□A) would be considered incongruent. Introducing actions that consist of two components (see Figure 2) renders the definition somewhat more complicated. Our participants were exposed to two pairs of actions and action effects during acquisition: A□1+C□3 and B□2+D□4. In the test phase, we presented the action effect of one of the second components (3 or 4) and we tested whether this would affect processes related to the first components. Hereby, the pairings of effect 3 and component A (3□A) or of effect 4 and component B (4□B) were considered congruent, and the pairings of effect 3 and component B (3□B) or of effect 4 and component A (4□A) incongruent. If infants represent the experienced action sequences as a unity, cueing the effect of the second element (C or D) could affect the activation of the first element (A or B). Finding any difference depending on congruency would provide evidence for a cognitive representation of sequential action in infants. Moreover, the direction of the effect would speak to the internal structure of that representation: While chaining and integrative models would lead one to expect facilitation (shorter latencies and smaller pupil dilations) in congruent responses, concurrent activation models would predict the opposite.

Methods

Subjects

Two age groups were tested: 14 9-month-olds (mean: 8.94 months, *SD*= .37, *SE*=.9, 5 female) and 16 12-month-olds (mean: 11.99 months, *SD*= .42, *SE*=.10, 9 female), another 4 9- and 7 12-month-olds were excluded for not meeting the criterion for the minimal amount of test trials. They were recruited through the municipality and received small gifts as compensation. An informed consent and a questionnaire regarding general health and development were obtained. The infants were all healthy full-term and without pre- or perinatal complications.

Test environment and apparatus

During the experiment the infants sat in a specially designed, stimulus-poor booth on the lap of their caretaker, who was seated in front of the eye-tracker apparatus. The distance between eyes and apparatus was approximately 70 centimeters (the screen’s viewing angle was 34.1° by 21.8°). The behavior of the infants was monitored online by the experimenter from a separate control room by means of a camera located above the apparatus. A 17 inch TFT-screen (1280 x 1024 pixels), equipped with an integrated Tobii T120 eye-tracker operating at 60 Hz, was used for visual and auditory data presentation, and data collection. The Tobii T120 has an average accuracy of .5° and allows for a certain amount of head movement by the subjects (30x22x30cm). It recorded gaze direction and pupil-size. Stimulus presentation was controlled by a PC running E-prime® software (Schneider et al., 2002).

Stimuli

The visual stimuli used were as follows (see Figure 2). The background color of the screen was grey. The fixation point was a brightly colored dot with a superimposed line drawing (4.3° by 4.3°). To keep infants interested, the color of the dot changed randomly from trial to trial (selected from eight colors) and the line drawing was randomly selected (without replacement) from a selection of 50 drawings (Snodgrass & Vanderwart, 1980). As Response Areas (RA's), we used 100 grayscale pictures from the "Nottingham scans" faces database, (http://pics.psych.stir.ac.uk), displaying emotionally neutral frontal faces of 50 men and 50 women. Faces were chosen to elicit spontaneous saccades as they are known to attract infants' attention (Goren at al., 1975; Johnson et al., 1991). To maximize the chance of finding an effect, the faces looked at the participant, since Sato and Itakura (2013) showed that eye contact enhances action-effect binding. We used two pairs of 200ms effect sounds which were equalized on loudness, "tring" and "piew" (Verschoor et al., 2013, 2012) and complex high- and low- note sound waves of 1574- and 776-Hz.

Procedure

Infants were tested at a time when they were likely to be alert. Prior to the experiment the caretakers were instructed not to move after calibration and gently hold the infant in order to maintain eye-tracker alignment, and to entertain the infant during the 1-min interruption between calibration and the experiment. The eye-tracker was calibrated using a 9-point calibration consisting of a small animation. The calibration was accepted with a minimum of eight points acquired. The experimenter could play an attention-grabbing sound during the experiment. If this no longer worked caretakers were encouraged to direct the infant's attention to the middle of the screen by pointing. Lighting conditions were kept constant. Furthermore, luminance levels were controlled for by presenting the stimuli in a random fashion. After completion an explanation of the experiment was provided.

Acquisition phase

The experiment began with an acquisition-phase of 36 trials (see Figure 2). If during the acquisition phase the subject showed declining attention, the acquisition phase could be shortened (minimum number of acquisition trials was set at 24). In each trial participants could freely choose to perform one of two saccade sequences. Each saccade sequence consisted of two distinct actions, first one to the left or right whereafter an up- or downward action followed (depending on the mapping assigned). Each saccade was followed by an effect-sound which was consistently designated to left-, right-, up- and downward Response Areas (RA's).

A trial started with the fixation dot. The dot disappeared after fixation on it for an interval that varied (to remove any bias or habituation caused by fixed intervals), between 150- and 350-ms. After disappearance, photographs of two different faces (randomly selected without replacement from 100 pictures) appeared to the left and right. The faces served as Response Area's (RA's). The 5.3° by 5.3° pictures appeared at 9.7° to center. To avoid perseverance to either side across trials the images pulsated. One of them started shrinking to 4.1° while the other started growing to 6.5° (side shrinking was randomized); one cycle from intermediate size to small, to intermediate, to large and back to intermediate, took 2 s.

When a saccade towards one of the faces was detected it stopped pulsating and the other face disappeared. Depending on the targeted side, one of two distinct 200ms effect-sounds ("tring" or "piew") was presented (the mapping was balanced across participants). RA's were defined as the maximum size of the pulsating images: 6.5° by 6.5°. A saccadic response was defined as eye movement (minimally 4.3°) into the left or right response area. Immediately after the effect the current face disappeared and reappeared 7.8° above or below that location (depending on the mapping) in the same dimension and continued to pulsate serving as RA again (again defined as the maximum size of the image). Upon detection of a saccade to that location (minimal 1.3°), one of two distinct 200ms effect-sounds ("high note" or "low note") was presented (the mapping was balanced across participants). RTs were defined as the time interval between disappearance of the fixation dot and detection of a saccade in the secondary RA. The maximum allowable RT was 2000ms; if by then no response was detected, the trial was repeated. After each trial, an inter-trial-interval of 500ms was used.

Test phase

The test phase of 32 trials followed directly afterwards (see Figure 2). The minimum number of test trials to enter analysis was 22. A trial started with the fixation dot as during acquisition. However, after fixation (fixation time identical to acquisition), the dot remained on display for 200ms during which an effect-sound was presented that was previously triggered by one of the two secondary eye-movements. Thereafter the dot disappeared. Then, two identical 5.3° by 5.3° images of the same face (randomly selected without replacement) appeared 9.7° to the left and right of the screen center serving as RAs. The two images were identical to minimize gaze preference. To further reduce bias the faces pulsated in synchrony, meaning that they either both grew or shrank (randomized and with the same motion parameters as during acquisition). Again, the images were expected to evoke saccades. The question of interest was whether the direction of these saccades would be biased by the tones. Except for absence of auditory effects after the saccades, the remaining procedure was as during acquisition.

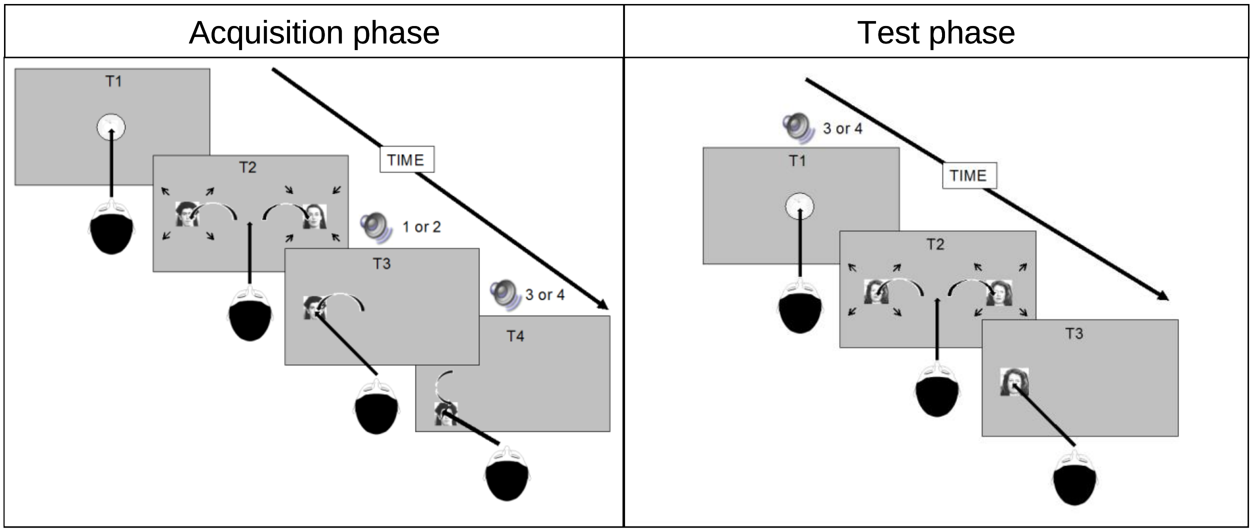

Figure 2. **Acquisition trial**: Each trial started with an intertrial interval of 500ms. T1: A fixation dot was displayed at screen center. T2: After successful fixation, faces appeared at either side of the screen where they started to pulsate. T3: Depending on the saccade target, the face at the other side disappeared and an effect sound was played for 200ms. T4: Depending on which side was chosen the face moved up or down. **Test trial**: Each trial started with an intertrial interval of 500ms. T1: A fixation dot was displayed at screen center. After succesful fixation one of the previous action effects was played. T2: The dot diasppeared whereafter the same face appeared on both sides. T3: The participant freely chose where to saccade.

Data acquisition

E-prime® 1.2 (Psychology Software Tools, Sharpsburg, PA) was used to collect RTs, the number of left and right responses and congruent and incongruent responses during test. The gaze- data files Tobii produced were imported into BrainVision Analyzer 1.05 (BrainProducts GmbH, Gilching, Germany) to analyse gaze position and pupillary data. Depending on analysis, segments were created from 2000ms before the presentation of the sound onset or RT, to 8000ms after, while allowing overlapping segments. Responses were sorted on congruency of the response and stimulus- and response-locked functions were averaged (Verschoor et al., 2013). Following Bradley et al. (2008), pupil-diameter measurement began after the initial pupil reflex caused by the fixation stimulus. Visual inspection showed it to end around 500ms after effect presentation (see Figure 4) (see also Verschoor et al., 2013). Dilations were calculated as the percentage of dilation relative to the baseline to make the data comparable across age groups. The percentage of trials rejected due to erroneous data points (leaving 29 valid trials on average) did not differ across age groups, p > .8. Dilations of both eyes were averaged to reduce noise. Artifacts and blinks detected by the eye-tracker were corrected using a linear interpolation algorithm, after which a 10 Hz low-pass filter was applied (c.f., Hupe, Lamirel & Lorenceau, 2009). Further artifact rejection was done using a threshold based approach, including those segments with pupil sizes between 1 and 5 mm, and a maximum change in pupil size of .03mm in 17ms. Gaze data were recorded in pixel coordinates, averaged between eyes and filtered using a 10Hz low-pass filter.

Given that the acquisition of action-effect associations is sensitive to the same factors as stimulus-response learning (Elsner & Hommel, 2004), the Number Of Completed Acquisition Trials (NOCAT) was taken as an individual measure of action-effect learning. The Mean Acquisition Reaction Time (MART) was taken as an individual measure for general speed and activity. Both NOCAT and MART variables were used as covariates in the analyses when appropriate.

Results and Discussion

Acquisition phase

First we tested for age group differences in dependent variables collected during acquisition to ensure that the learning experiences of the age groups were comparable (see Table 1). All ANOVA’s were performed with age group as a between-subjects factor. There were no effects for the percentage of completed acquisition trials (p >.5), mean RT (p >.5), or the percentage of right vs. left responses (p >.2) or upward vs. downward responses (p >.2). Two reliable effects were obtained for RTs. Firstly, horizontal response location interacted with age group, *F*(1,25)=6.25, *p*= .019, η2p=.20. Separate analyses showed no main effect in 9-month-olds (RT-left=999ms, RT-right=104 6ms) and a tendency toward faster rightward responses in 12-month-olds, *F*(1,12)=3.84, *p*=.074, η2p=.24 (RT-right=982ms, RT-left=1089ms). Secondly, vertical response location interacted with age group, *F*(1,25)=4.63, *p*=.04, η2p=.16. Separate analyses showed no main effect in 9-month-olds (RT-up=1008ms, RT-down=1037ms) and a tendency toward faster downward responses in 12-month-olds, *F*(1,12)=3.82, *p*=.07, η2p=.24 (RT-up=1089, RT-left=982ms). We also performed a repeated measures ANOVA on RT's with Time (dividing the responses in three equal bins) and found no effect (*p* >.13). Lastly, we performed a repeated measures ANOVA on the partial RTs of the primary action to test if contraction vs. expansion had an effect on these partial RTs. We found a significant effect, *F*(1,28)=175, *p* <.001, η2p=.86, indicating responses toward contracting pictures were slower (partial RT-contracting=603ms, partial RT-expanding=428ms).

*Table 1***.** Mean scores of acquisition phase (standard deviation in brackets).

AGE GROUP SCORES	Percentage of completed acquisition trials	Percentage of left responses	RT in ms	RT left	RT right	RT up	RT down
9-month-olds	92.9 (11)	42.8 (38)	1007 (81)	999 (107)	1046 (90)	1008 (116)	1037 (83)
12-month-olds	92.0 (14)	38.5 (40)	1023 (83)	1089 (167)	982 (73)	982 (78)	1089 (164)

We concluded that the learning experiences were comparable across age groups. The interaction of horizontal response location and age on RTs might reflect the fluctuating emergence of general right-side preference during the first year (Corbetta & Thelen 1999; Michel, 1998), which also affects infants' eye movements (Cohen, 1972). An orthogonal effect may be reflected in our analysis of upward vs. downward RTs. However, little is known about such preferences. Additionally, we found that infants responded faster toward expanding pictures. This effect probably reflects automatic attentional processes to avoid collisions (e.g., Kayed & Van der Meer, 2000; Van Hof, Van der Kamp & Savelsbergh, 2006). Importantly, these observations are not detrimental to our research question since both age groups received approximately the same amount of training for all combinations of response locations.

Test phase

All ANOVA's were performed with age group as a between-subjects factor. There was no effect on the percentage of completed test trials ($p > .4$).

Response frequency

Overall, participants looked more often (64%) to the right than left side, $F(1,28)=9.00$, $p=.02$, $\eta 2p=.19$, but the effect did not interact with age. More important for our purposes, ANOVA's with congruency as within-subjects factor were not significant, adding MART, NOCAT or both as covariates didn't change this (*p*'s >.2). We concluded that, if infants control sequential actions, this does not seem to affect the probability to choose a particular sequence.

Reaction times

There were no reliable effects with regard to overall RT ($p > .6$), left vs. right response location ($p > .3$) (see Table 3) or inter-trial interval, ($p > .5$) (which we analyzed because the test-phase was self-paced). More important for our purpose, an ANOVA with congruency as within-subjects factor, revealed a significant effect indicating 29ms-slower responses for congruent trials, *F*(1,28)=4.15, *p*=.05, η2p=.13; the interaction with age was not significant (*p* >.3). Although the statistics did not necessitate further exploration, given our directed hypothesis about age effects, we looked at both age groups separately. In the 9-month-olds the effect was not significant (p >.4) while in the 12-month-olds it was *F*(1,15)=5.47, *p* =.03, η2p=.27. A non-parametric Wilcoxon signed rank test confirmed these results (9-month-olds: Z=-1.57, p=0.88, 6 of 14 infants showed the pattern, 12-month-olds: *Z*=-2.02, *p*=0.04, 12 of 16 infants showed the pattern). However, adding NOCAT as a covariate into the separate ANOVA for the 9-month-olds resulted in a significant effect (*F*(1,12)=4.96, *p*=.05, η2p=.29).

*Table 2***.** Mean frequency and RT scores of test phase (standard deviation in brackets).

AGEGROUP SCORES	Percentage completed test trials	Percentage left responses	Percentage congruent responses	ITI ms	RT ms	RT Congruent ms	RT Incongruent ms
9-month-olds	93.5 (11)	43.2 (38)	47.5 (8)	1637 (323)	431 (90)	440 (104)	424 (91)
12-month-olds	96.5 (10)	29.0 (21)	49.3 (7)	1563 (393)	447 (74)	468 (83)	425 (83)

Chapter 8

Our main finding is: cueing of the secondary element of the action sequence interfered with executing its first, as evidenced by the longer RT's for congruent responses in the 12-month-olds. Results were less clear in the 9-month-olds. However, using NOCAT as a covariate resulted in a similar effect in 9-month-olds, suggesting that the Number of Completed Acquisition Trials was an important factor for the strength of the effect in this age group. The fact that we found an effect can be considered as evidence that 12-month-olds control sequential action. Performing two consecutive actions is sufficient to integrate them into a coherent representation. Twelve-month-olds apparently represent action sequences in a format that allows for interactions between the codes of their individual elements (which excludes fully symbolic formats). Moreover, our findings provide specific support for concurrent activation theories, as only these would predict interference. Furthermore our findings suggest sequential action control is developing in 9-month-olds.

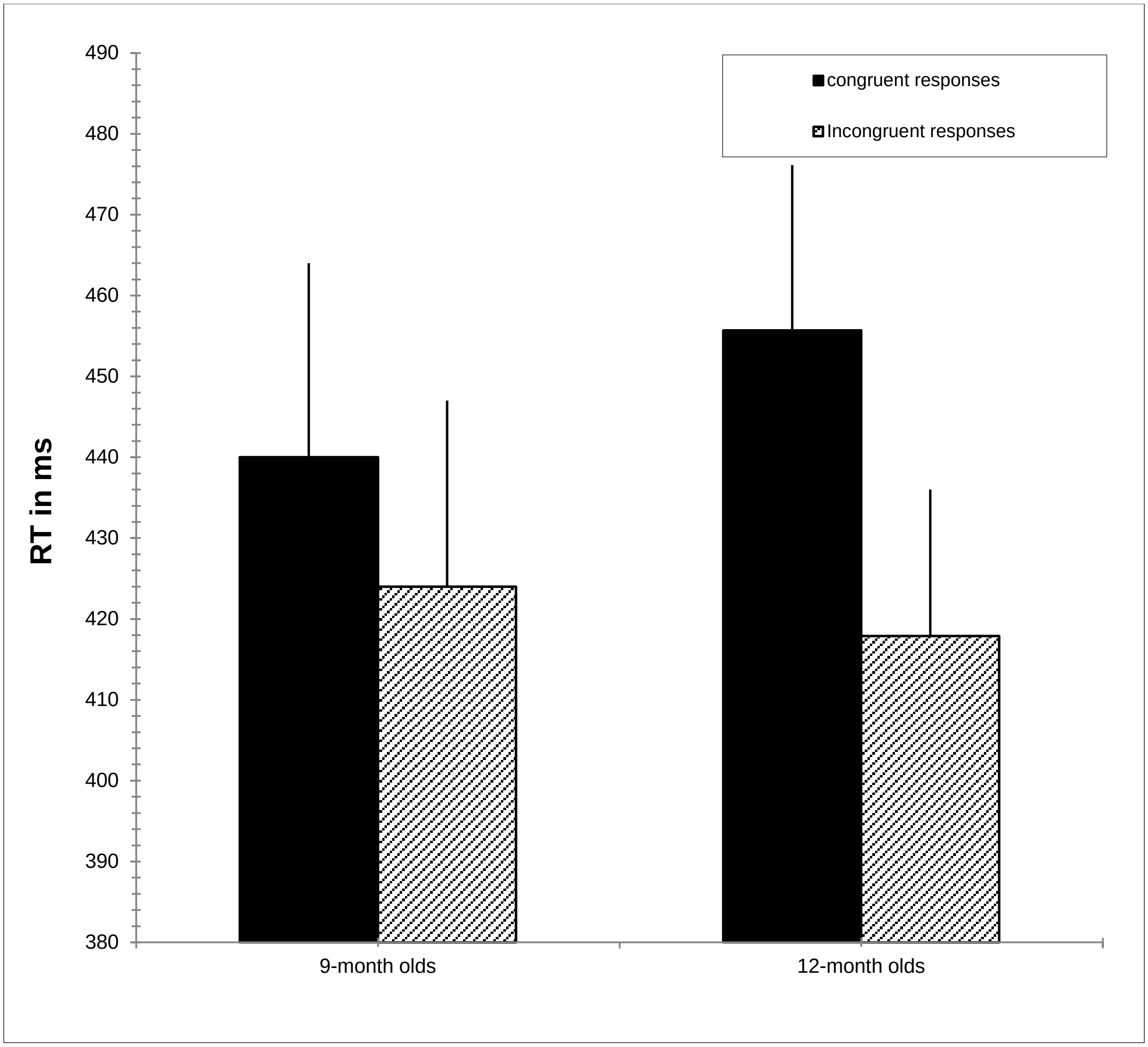

Figure 3. Mean reaction times (+SE) for 9-month-olds (N=14) and 12-month-olds (N = 16) in congruent and incongruent test trials.

Pupil dilation

To accommodate for the variable RTs across age groups and conditions, we considered both stimulus-locked and response-locked Task-Evoked Pupillary Responses (TEPR's). The stimulus-locked analysis for congruent and incongruent responses TERPs used a 500ms pre-effect baseline (Beatty & Lucero-Wagoner, 2000). A repeated measures ANOVA on TERPs with congruency as within subjects factor revealed no a priori effects of congruency on baselines (-500 to 0ms), *p*'s > .7. Adults' TEPRs start from 200 to 300ms after stimulus onset and peak in the range of 500ms to 2000ms (Beatty, 1982; Beatty & Lucero-Wagoner, 2000). We therefore calculated the mean TERPs for congruent and incongruent responses as the mean percentage of change from baseline to 500-2000ms post effect onset. An ANOVA with MART as covariate revealed that, overall, participants exhibited larger relative pupil dilations during congruent responses, *F*(1,27)=4.12, *p*=.05, η2p=.13, independent of age group (*p* >.7). Since the time window was based on adult findings, which likely underestimate the pupillary reactions of the slower infants (Verschoor et al., 2013), we reran the analysis with a 1000-2500ms post effect onset time window. Again, pupil dilations were significantly larger in congruent trials, *F*(1,25)=5.03, *p*=.03, η2p=.16, independently of age, *p* >.09 (see Figure 4).

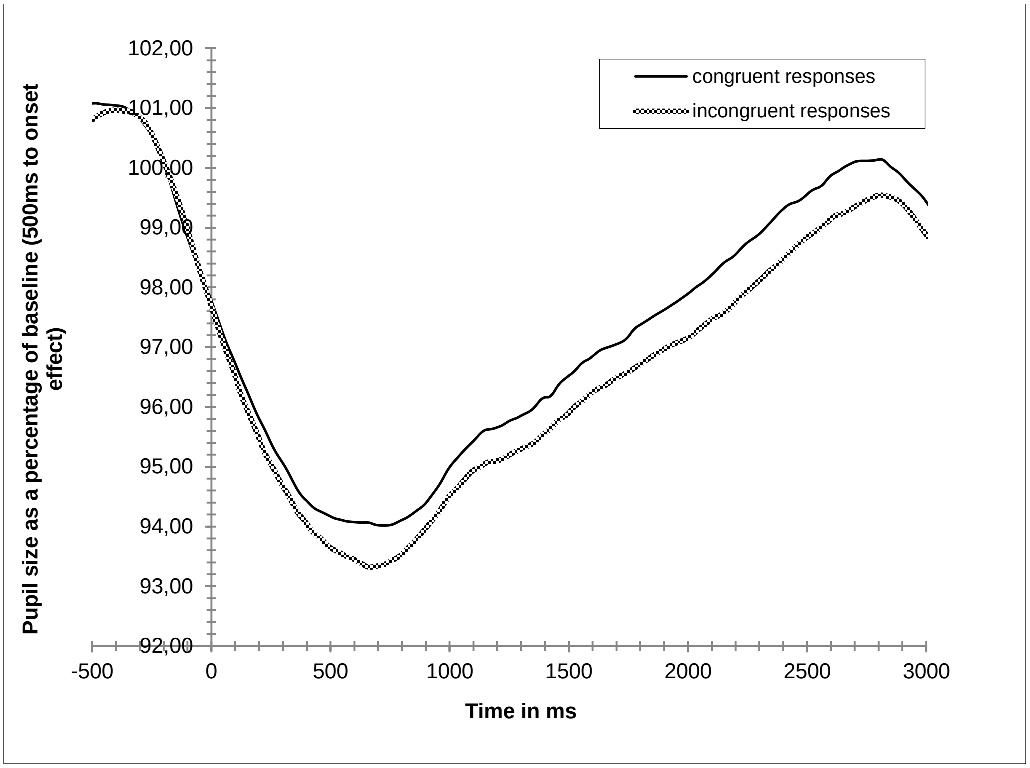

Figure 4. Relative pupil sizes for congruent and incongruent responses to baseline, stimulus-locked.

For the response-locked analysis, we calculated the percentage of dilation from a 700-ms time window starting at saccade onset, to the same 500ms pre-stimulus baseline. An ANOVA with MART as covariate yielded a tendency for larger relative dilation in congruent trials, *F*(1,27)=3.51, *p*=.07, η2p=.12, while the interaction with age group was not significant, *p* >.8 (see Figure 5). Adding NOCAT as additional covariate resulted in a significant effect (*F*(1,26)=5.48, *p*=.03, η2p=.17), again without an interaction with age group (*p* >.9).

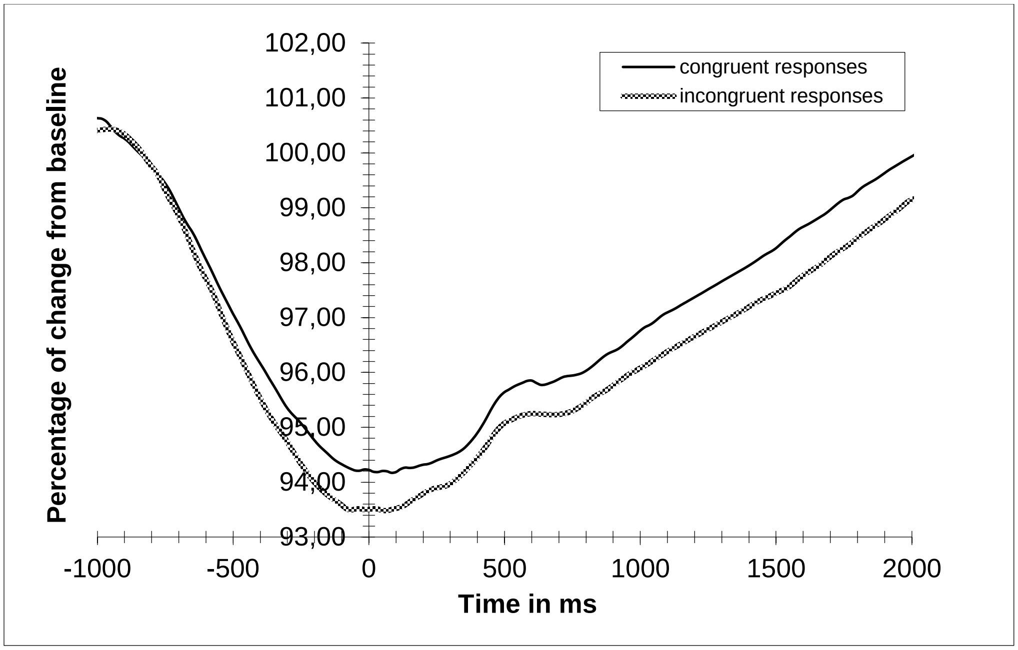

Figure 5. Relative pupil sizes for congruent and incongruent responses to baseline, response locked.

Finding larger relative pupil dilations for congruent responses in both stimulus-locked and response-locked analyses corresponds nicely to the outcome of the RT analysis. Cueing the second component of an action sequence makes the execution of the first slower and more effortful.

Gaze position

In the test phase we primed the second element of the two-element sequences carried out in the acquisition phase by presenting the corresponding action effect. Activating the second element of the sequence might affect action planning directly. One of the second elements was an upward movement while the other was a downward movement. Priming these elements by their effects might induce a vertical bias in the direction of the cued element. Alternatively: the selection of the primary action results in forward inhibition of the second. This might induce the opposite bias. To investigate these effects, we analyzed the mean vertical deviation from the horizontal midline toward the primed action element as a function of congruency. To do this we collapsed all vertical deviations from horizontal midline toward the direction cued to one side and divided the data segments according to congruency from stimulus onset to 650ms thereafter (corresponding to the mean RT plus mean random ITI) and compared these segments to a 150ms pre-effect baseline (the minimum fixation time before effect onset).

There were no a priori effects of congruency on baselines, $p > .5$. An ANOVA with MART as covariate showed that during congruent responses gaze position deviated vertically significantly less toward the direction cued by the effect sound, $F(1,27)=4.83$, $p=.04$, η2p=.15 (effect size = 22 pixels; see Figure 6) than in incongruent responses, and this effect did not vary with age, $p > .5$. We additionally performed separate ANOVAs with MART as covariate testing congruent- and incongruent- responses against no deviation. The effect was significant for incongruent responses ($F(1,27)=4.70$, $p=.04$, η2p=.15) and did not vary with age ($p=.2$), but not significant for congruent responses ($p > .26$).

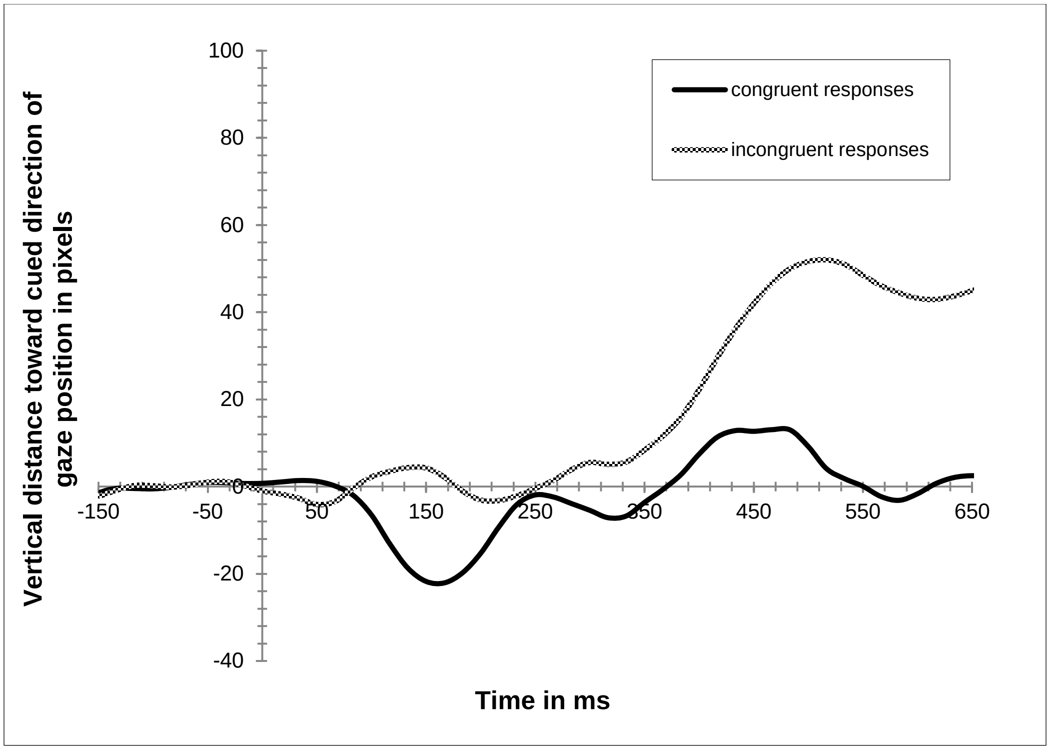

Figure 6. Vertical distance from midline toward cued direction of gaze position in pixels for congruent and incongruent responses. Time=0, is the moment the effect starts.

Our gaze position findings show that priming the second action component results in activation of the vertical component only in incongruent trials. This might be due to competition between activated components in congruent trials, as concurrent models would hold: the selection of the primary action results in forward inhibition of the second. This finding provides further evidence for the concurrent model of sequential action control in infants and highlights its temporal dynamics. Moreover, the finding is not in accordance with integrative models since no vertical bias was found for the primary actions in congruent trials.

General discussion

The aim of the current study was to examine (the development of) the cognitive substrate for sequential action control in 9- to 12-month-olds using a novel gaze-contingent paradigm. Relying on the idea that if two elementary actions are bound together in a sequence, priming the secondary action component would influence the availability of the primary component, we presented the infant participants with a two-step action sequence. While chaining and integrative models would lead one to expect facilitation in congruent responses, concurrent activation models would predict the opposite. Our major finding is that priming the second action inhibits the primary action, as indicated by latencies and pupil dilation. Secondly, we found an effect on gaze position indicating that action control inhibits the second component of an action sequence whilst preparing the first part of the sequence. Our findings on three different measures suggest an emerging ability for sequential action control in 9-month-olds that fully develops by the first birthday, and is best captured by concurrent activation models (Estes, 1972).

From a developmental perspective our findings extend behavioral studies suggesting infants can control sequential action (e.g., Claxton et al., 2003; McCarty et al., 1999), and studies showing that this ability to be present only under ideal circumstances in 9-month-olds, or only a subset of subjects of this age group (Bauer, Wiebe, Waters & Bangston, 2001; Carver & Bauer, 1999, 2001; Elsner et al., 2007; Lukowski et al., 2005 Waters & Bangston, 2001). In addition, they are in accordance to studies suggesting that during the second half of the first year the ability to encode ordinal information comes online (Brannon, 2002; Picozzi et al., 2010; Suanda et al., 2008). The fact that evaluation of third-person sequential action is apparent significantly earlier in development, in 6- to 7-month-olds (Baillargeon et al., 1990; Biro et al., 2011; Csibra, 2008; Gergely & Csibra, 2003; Verschoor & Biro, 2012), tentatively suggests either a different cognitive substrate or a dissociation between evaluation and production (e.g., Verschoor et al., 2013).

Furthermore, our findings relate to infant studies (Claxton et al., 2003; McCarty et al., 1999; Cox & Smitsman, 2006) and cognitive theories (e.g. Botvinick & Plaut, 2004; Constantinidi et al., 2002; Cooper & Shallice, 2006; Estes, 1972; Norman & Shallice, 1986 Rumelhart & Norman, 1982) that implicate inhibition as an imperative faculty for controlling sequential action. Interestingly, inhibitory control begins to emerge toward the end of the first year and undergoes rapid development across the toddler period and into the preschool years, a pattern coinciding with age-related changes in frontal lobe maturation and connectivity (Diamond, 2002, Diamond et al., 2007, Luria, 1973; Wolfe & Bell, 2007). This onset around 1 year of age relates to the developmental timeline revealed by our results and supports our interpretation that inhibitory processes play an important role in the ontogenesis of sequential action control.

Note that the development of inhibitory capacities has been linked to the development of time perception itself (Zélanti & Droit-Volet, 2011; Mäntylä, Carelli & Forman, 2007). Furthermore, in clinical (Barkley, 1997; Gerbing, Ahadi & Patton, 1987; Montare, 1977) and healthy populations (Foster et al., 2013) tests of inhibition show robust relationships to indices of timing (-deficiency). Thus the question arises whether the inhibitory mechanisms found are specific for action control or are general for representing temporal events (Fuster, 1993, 2002; Norman & Shallice, 1986). The literature reviewed here seems to point to the latter, suggesting temporal representations in the form of concurrent activation may be a precondition for sequential-action control. One might thus speculate that very early sequential-action evaluation (e.g., Verschoor & Biro, 2012) depends on non-ordinal, or non-temporal representations. Interestingly, our paradigm offers a possibility to address these and related questions in future research.

Concerning current theories on action control, our results seem to point to limitations in explanatory power of the current ideomotor theory (Hommel et al., 2001; Shin, Proctor & Capaldi, 2010) with regard to sequential action control, as this theory would predict activation of actions by their effects whereas we find inhibition of the primary action. Sidestepping the idea that inhibition of the primary action is not the same as inhibition of the sequence as a whole, Hull (1931) pointed out that binding of sequential action is possible by relating the sequences to internal or external context such as an overarching goal. An interesting question that such reasoning poses, is what kind of context and how such context may be incorporated in an overarching goal representation. Indeed action-effect learning can be context-specific (Kiesel & Hoffmann, 2004) which could accommodate such overarching ideomotor representations of action sequences. Thus we may not have succeeded in cueing the overarching goal because of insufficient context in the cue, and might have gotten stuck in the underlying concurrent level of representation. Indeed, Kiesel and Hoffmann (2004) have shown that the same actions can be accessed by different effect anticipations. They also claim that that response initiation has to wait for the anticipation of the effects that trigger the response (see also Kunde, 2003), suggesting it takes longer to initiate a response if it produces a long effect. Although this theory would also predict slower initiation for sequential actions, if one thinks of a sequence of (actions and) effects as a long effect, the theory cannot account for competitive processes our findings suggest. Thus we suggest ideomotor theory should be enhanced by incorporating overarching- and sequential- levels of goal representation. Such hierarchical structure might be conceived as either structural (Cooper & Shallice, 2006) or epiphenomenal (Botvinick & Plaut, 2004) to concurrent activation models.

Nonetheless, our findings do suggest ideomotor processes play a role in sequential action since we inhibited the primary and secondary action components by cueing the secondary action via its effect. Ideomotor processes have indeed been implicated in sequential action (Koch, Keller & Prinz, 2004). Ziessler (1994, 1998) and Elsner et al. (2007) found that action-effects play an important role in sequence learning and Stöcker & Hoffmann (2004) found that action effects facilitate chunking. Furthermore, the activation of the secondary vertical component in the incongruent trails is direct evidence for ideomotor theory. Thus although the current findings extend cognitive theories of action control by suggesting that ideomotor theory needs elaborations to incorporate sequential action, they do not counter the ideomotor principle itself.

Another theoretical implication of our results is that the repeated successiveness of actions in the acquisition phase sufficed to bind the actions into a sequence. This raises the question of what the exact criteria might be for such binding to occur. One could think of several dimensions for such criteria; our study suggests repetition, temporal closeness and spatial closeness might play a role. This is a particularly interesting question since its answer might provide clues as to how the cognitive system generates new action sequences, never performed before. However, more research is needed to answer such questions in more detail. For now, our results suggest that concerning infants' own action control, sequential action can be picked up by exploration and does not necessarily depend on elaborate abstract or explicit strategies (Cleeremans & McClellend, 1991) that operate in terms of efficiency (e.g. Gergely & Csibra, 2003) or causality (e.g., Woodward & Sommerville, 2000).

Even though we consider the present findings as a first step towards the understanding of sequential-action control in infants, further research is needed to explore this model in greater detail. Although our paradigm produced continuous data which are temporally rich, they nevertheless should be considered as a snapshot of processes at work in sequential-action control. Earlier we hypothesized that cueing the secondary effect of the two-step sequence might have been too poor in contextual information to cue the overall sequence, thus resulting in local competition effects in an underlying concurrent level of representation. Alternatively, more dynamic explanations could be considered. For instance, it could be that the inhibition of the first sequence components was due to temporal differences in the process of activating the individual action components on the one hand and of the overarching goal representation on the other. It is well documented that initiating more complex sequential actions takes longer than initiating simpler actions (Henry & Rogers, 1960; Rosenbaum, 1987). One could thus speculate that cueing the second action component activated the underlying representation quickly, but it took more time to activate the overarching goal representation. The eventual activation of this goal representation could have facilitated both components of the sequence (as proponents of integration theories might suggest), but that may have taken too long to be picked up by our measures. As a consequence, the inhibition that our findings point to may reflect an initial state of a dynamic action-planning process. Another possibility would be that cueing an action component that is not yet appropriate (as none of the secondary components was a valid action in the test phase) resulted in the inhibition of not only the first component but of the entire sequence, perhaps including the goal representation. We cannot exclude that the second component of each sequence was also inhibited—although the lack of gazing "away" from the direction cued by the second component suggests that it was not. The current experiment was not set up to distinguish between these more detailed scenarios.

Other studies have shown end state comfort effects in infants (Claxton et al., 2003; McCarty et al., 1999; Cox & Smitsman, 2006) indicative of integrated representations of sequential action. In the current study we did not find evidence for this model. Nonetheless, we do not wish to claim that integrated sequential action control cannot occur in infancy. We would like to stress that the chaining, concurrent and integrated theories of sequential-action control are by no means mutually exclusive or complete. They posit useful approximations for understanding sequential actions, yet depending on exact circumstances relating to practice, content, time pressure and strategy, some models may be more adept than others at explaining specific empirical phenomenon. In our opinion a future all-encompassing theory of sequential action control will probably encompass elements of all three classes of theories. Indeed our results on gaze directions indirectly suggest an influence of the secondary action on the primary action that was cancelled out by counteracting inhibitory and excitatory processes. However, this does not diminish the importance of showing that concurrent processes are at work in infant sequential-action control.

It remains essential to develop such a comprehensive theory of (the development of) sequential action representation, which specifically addresses the question of how novel components are integrated into a sequential plan, how the sequence is generated, whether this requires hierarchical representations and what types of information are incorporated in overarching goals. We are confident that the present paradigm can be helpful in answering some of these questions by introducing further modifications of the task (e.g., by cueing the first component and examining what effect this would have on the availability of the second). Since saccade-effect learning is now well established in adults (Huestegge & Kreutzfeldt, 2012; Herwig & Horstmann, 2011) and infants (Verschoor et al., 2013) we are confident that the paradigm can help to increase our insight into general cognitive mechanisms underlying action planning since our synergy of methodology provides various measures (frequency-, RT-, pupillary- and gaze position measures) that can pick up different dynamic aspects of the planning process.

In conclusion, the current study shows that sequential action can be picked up by exploration and does not depend on elaborate abstract strategies. Furthermore, the present findings demonstrate that 12-month-olds are able to construct action plans comprising more than one element, and use inhibition mechanisms as suggested by concurrent activation models to put elements into the right temporal order. And lastly, we provide further evidence for the claim that the ability for sequential-action control develops between 9 and 12 months of age.

Chapter 9: Summary and conclusions

Summary and Epilogue

The aim of the current thesis was to answer some of the questions which current developmental (neuro-) cognitive theories pose with regard to how infants learn to perceive, and perform goal-directed action. There is consensus among scholars that during the first year of life infants become sensitive to the goals of observed actions (e.g., Woodward, 1998), to the means by which they are achieved (e.g., Gergely et al. 1995) and use this information to predict future events and guide their own action (e.g., Elsner & Aschersleben, 2003). Nevertheless, the cognitive mechanisms that subserve these abilities are still under debate. Before discussing in more detail the (preliminary) answers the current empirical work offers, the main findings of the empirical chapters of the thesis will be summarized. Lastly, the implications for further research and my current *opinion* on what theoretical view best captures the current findings in the field will be presented.

Summary

Chapter 2 presents a study that investigated the relative importance of outcome-selection information vs. means-selection information. Many studies that use Woodward's (1998) VoE paradigm show that outcome-selection information (information presented by the choice an actor makes between potential outcomes) is important for infants' goal perception. On the other hand, a significant number of studies in the tradition of Gergely et al. (1995) show that means-selection information (information presented by the efficiency of the action toward the outcome in regard of the situational constraints) is important for infants' goal perception. Until now the relative importance of these two types of information has not been directly compared, partially due to the conceptually different paradigms used to investigate them. In this study these two types of paradigms were combined. It was found that when outcome selection information was presented, but the means were inefficient toward the goal, 7- and 9-month-old infants did not attribute a goal to an observed action. This finding suggests that means selection information displays primacy over outcome selection information not only for adults but also for infants. The early presence of this bias sheds light on the nature of the notion of goal in action understanding. Furthermore, the central claim of the “theory of rational action” - that infants *only* attribute a goal to an action if the action can be evaluated as an efficient action toward the end-state in the given situation - has so far only been supported by relatively indirect evidence. The experimental paradigm that had been used to investigate this theory tested infants’ predictions about the means, and not the goal, of the action. However, the finding that infants do not generate a specific expectation about the new means if the initial action is non-efficient only suggests, but does not prove, that infants do not in the meantime attribute a goal to the initial action. Infants may have given up on predicting the means, but could in principle have relied on the end-state of the action to infer it as the goal. The current paradigm enabled testing the expectations about the goal of the action by providing two possible end-states. The finding thus confirms that if the action, the end-state, and the situational constraints do not form a relation that satisfies the principle of efficient action, then infants do not commit themselves to a specific goal. In sum, this study showed that infants' early understanding of goal-directed actions is similar to that of adults as far as the relationship between means selection and outcome selection information is concerned. The early presence of the preferential bias for means selection information may suggests that this bias is not an acquired, but rather a core property of the cognitive mechanisms of goal attribution and thus it sheds light on the nature of the notion of goal.

Chapter 3 utilizes a similar paradigm to chapter 2, integrating the Woodward- (1998) and Gergely et al.- (1995) type paradigms to further investigate how means-selection information relates to outcome-selection information. It is possible that infants appeal to one goal-like notion (e.g., 'preference') in situations that provide outcome selection information, and to another one (e.g., 'planning') when they receive evidence on means selection. Alternatively, a unitary concept of goal explains the results of both kinds of studies, which takes input from either type of information. In the latter case, one would expect transfer of goal attribution from situations with one type of information to situations in which the other type of information is available. Thus the central question in this chapter is if infants can transfer goal attribution from a situation in which only means information is presented toward one goal, to a situation in which there two possible outcomes. In other words, does observing an actor efficiently adjusting his action to situational constraints, lead infants to expect that the actor will continue to approach the same goal after another potential outcome is presented in the scene? Such transfer would indicate that both the Woodward- (1998) and Gergely et al.-(1995) type paradigms tap into the same unitary concept of goal that can rely on both outcome- and means- information. It was found that 12-month-olds who had attributed a goal based on the causal efficiency of a means-end action, generated expectations about the actor's action in another scenario in which the actor could choose between alternative outcomes. This finding thus suggests that, by 12 months, infants possess a unitary concept of goal.

The fourth chapter explores the question how experience with a certain movement can influence goal attribution. More specifically, the hypothesis was investigated that, linking a novel first-person and/or third-person action to a salient action effect, is critical for infants to interpret a novel third-person action as goal-directed. The findings suggest that 12- but not 9-month-olds, provided they have previously associated the novel action with a salient visible outcome in another context, can assign a goal to the action even in the absence its outcome. On the other hand, the control condition suggests that prior experience with the action, but without the salient effect, does not lead to goal-directed interpretation of the novel action. The finding thus demonstrates the essential role action effects play in the developing ability of infants’ goal-directed understanding of novel actions.

To further validate ideomotor theory as a cognitive developmental theory, chapter five replicates previous findings (Eenshuistra et al., 2004; Kray et al., 2006) that indicate that 4-year-old children can form bidirectional action-effect associations and compares their performance to adults. The secondary aim of the study, was to investigated whether action-effect associations are also acquired under explicit-learning conditions, and whether familiar action-effect relations (such as between a trumpet and a trumpet sound) are learned the same way as novel, arbitrary relations are. Findings suggest that explicit learning produces the same bidirectional action-effect associations as implicit learning does, that non-arbitrary relations improve performance without affecting learning per se, and adults and young children show equivalent performance—apart from the common observation that children have greater difficulty to withstand stimulus-induced action tendencies.

The next two chapters represent the first attempts to provide direct evidence for the spontaneous acquisition of bidirectional action-effect associations in infancy. In chapter six a highly simplified version of the Elsner and Hommel paradigm (2001) was applied. First evidence for the spontaneous acquisition of bidirectional action-effect associations in 9- 12- and 18-month-olds was found, suggesting that the mechanism underlying action-effect integration is in place at the latest around 9-months-old.

Chapter seven is methodologically more advanced adaptation of the Elsner and Hommel paradigm (2001). The paradigm was made suitable for all age groups ranging from 7-month-old infants to adults, and employed a novel pupillometric and oculomotor paradigm to study developmental changes in the role of action-effects in the acquisition of voluntary action across the lifespan. The findings suggest that both 7- and 12-months olds (and adults) can use acquired action-effect bindings to predict action outcomes but only 12-months-olds (and adults) showed evidence for employing action-effects to select actions. This dissociation supports the idea that infants acquire action-effect knowledge before they have developed the cognitive machinery necessary to make use of that knowledge to perform intentional actions.

The last empirical chapter of the thesis used the newly developed methodology from chapter seven to investigate, how infants represent sequential action from an ideomotor point of view. The aim was to contrast chaining-, concurrent- and integrated models of sequential-action representation. Nine- and 12- month-olds were taught action sequences consisting of two elementary actions. Thereafter the secondary action was selectively activated to assess any interactions with the primary action. Results suggest that concurrent models best capture the representations formed.

Chapter 9

Discussion

Chapter 2, 3 and 4 were concerned with the question how infants learn to perceive others' actions. In general, the first three chapters of this thesis provide evidence that third-person action perception processing relies on several different sources of information. Chapter 2- and 3 show that the efficiency of an action is an important source of information for goal perception. Chapter 2 further suggests that it is more important than ends-selection information. Chapter 4 goes on to provide evidence for the notion that action-effect knowledge plays an important role in action perception.

Taken together the first three studies show that during the first year of life infants acquire a unitary concept of goal. Chapter 2 shows that at least by 7-months of age previously obtained means-selection information can transfer to predictions in situations involving two potential outcomes, and at least by 12-months of age, action-effect knowledge can transfer to predictions in situations involving two potential outcomes (Sommerville et al. (2005) provide evidence for similar transfer at 3.5-months of age). These transfers suggests that at the latest by age one, all three types of information feed into a unitary generalized concept of goal. Such a unitary concept of goal suggests a more sophisticated higher order interpretation of the current findings regarding infant goal attribution than just preference- efficiency- or action-effect detection. It suggests infants to recognize a "common denominator" in different aspects of human behavior, namely planning, action-effect relations and the expression of preference.

Nevertheless, in the current thesis both types of transfer: (1) from means selection information to situations that involve ends-selection information, and (2) from action effect knowledge to situations that involve ends-selection information, have only been shown in one direction. In principle, to truly show that these types of information are processed in reference to a unitary goal concept, transfer should also be possible in the opposite direction. Showing transfer in the opposite direction of (1) should be impossible. On the other hand, showing transfer in the opposite direction of (2) is harder to test since the selection of an end is to a certain extent the same thing as an action effect. Additionally, transfer from means selection information to situations where action effects are presented (3) and its reverse, should in theory also be possible. How would one go about to test these hypothesis, or is there evidence corroborating these hypothesis?

To test the first assumption (transfer in the opposite direction of (1)) one would need to design a new VoE paradigm. In the familiarization phase infants should see an actor repeatedly touching one of two toys. Where after, in the test phase infants should be shown the actor reaching efficiently or non-efficiently for the same toy (the other toy is not presented in the test phase). If transfer has occurred infants should expect an efficient approach to the toy and therefore should look longer at an inefficient action. On the other hand, if no transfer occurred they should not have a specific expectation as to the efficiency of the action shown. As a control condition infants should be shown efficient and inefficient reaches toward the new toy (the old toy is not presented). Since no goal attribution took place for this toy infants should not differentiate between the two test events. Further research should investigate this hypothesis.

The second proposed reversal of transfer (2) is already tested indirectly in 8-month-olds using EEG by Paulus et al. (2013). In their study they showed that the effects of actions that they had seen their parents do, elicited greater motor activation in infants than effects that were not associated with actions. This evidence indicates that infants can acquire action effect associations by observing others’ actions.

The third type of transfer (3) (transfer from means selection information to situations where action effects are presented) is again testable. One would need to adapt the Gergely et al. (1995) paradigm in the following way. The familiarization should be conceptually the same as in the original paradigm; a small ball repeatedly approaches a big ball. In one group of infants the small ball adjusts its approach to changing situational constraints with the only addition that the small ball starts in the middle. In the other group the ball approaches the big ball inefficiently, again starting from the middle. In the test phase the big ball is located on the opposite side as compared to the familiarization phase. The test phase should have two potential outcomes; one wherein the small ball approaches the big ball in a straight line and another wherein the small ball moves away from the big ball in a straight line. If transfer occurred, infants should expect the same outcome (or action effect) again (the big ball touching the small ball) even though it moves in the opposite direction as in the familiarization phase. Furthermore, this effect should only occur for the group of infants that were familiarized to the efficient approach. The opposite direction of the proposed transfer can also be tested. In the familiarization phase a small ball starting from the middle moves towards a big ball. In the test events the small ball moves toward the big ball which is now on the other side and adjusts its path to efficiently jump over an obstacle or the ball efficiently moves away from the bigger ball, efficiently jumping over an obstacle. If transfer took place, infants should look longer at the efficient jump away from the big ball.

Providing additional evidence for the reverse transfer of (1) and (2), and providing evidence for (3) and its reverse, would further strengthen the notion of an overarching unified concept of goal. As mentioned earlier such a unified concept would indicate a rather advanced goal perception. Furthermore, the developmental pathway of the integration of these types of information into a unified concept is of great interest to elucidate the ontology of the concept of goal, and should be investigated further.

The second part of the thesis was concerned with how actions are learned by experience. There is ample corroborating evidence suggesting (just as chapter 3 does) that infants use action-effect associations for third-person action perception (for a review, see: Hauf, 2007; Kiraly et al., 2003), imitation (for a review, see: Elsner 2007; Meltzoff, 2007) and exploration behavior (Elsner & Aschersleben, 2003). However, these studies do not provide direct evidence for the crucial bidirectional quality of action-effect associations as proposed by ideomotor theory. Many researchers nonetheless assume that learning new goals (learning new bidirectional action-effect associations) should be supported by the same mechanism as that found in adults (Elsner & Hommel, 2001). Thus the first aim of the second section of the thesis was to generate direct evidence for bidirectional action-effect uptake. The secondary aim was discover the developmental pathway the uptake of bidirectional action-effect associations follows.

After we succeeded in replicating the findings of Eenshuistra et al. (2004) and Kray et al. (2006) that indicate that 4-year-old children can form bidirectional action-effect associations similarly to adults, we went on to search for the same mechanism in infants. In chapter 6 and 7, we present first direct evidence for the uptake and use of bidirectional action-effect associations in infancy and therefore achieved our primary aim.

Furthermore, combining the findings of chapter 6 and 7, our findings suggests a major change in action-effect learning from just action monitoring to action selection, just before the ninth month of age. In chapter 6 we found evidence indicating that learned action-effects influence action control starting at 9 months of age and do so progressively more at 18 months of age. In chapter 7 we showed that although infants at 7 months of age do take up action-effect knowledge, they do not use it (under the conditions of the experiment) for action control. The dissociation we obtained in 7-month-olds suggests a developmental precedence of action monitoring over intentional action selection. Additionally we showed that 12-month-olds and adults showed employed action-effects to select actions in a similar fashion indicating that the same mechanism found in adults is already in place during infancy. Taken together the studies suggest a major change in action-effect learning from just action monitoring to action selection, just before the ninth month of age.

In chapter 8 our findings demonstrate that young infants are able to construct action plans comprising of more than one element and that they do so in a manner that puts the available elements into the right order. Even though we consider the findings as a first step towards the understanding of sequential-action representation in infants, the details of the suggested scenario are not entirely clear yet. We nonetheless present first evidence that infants are able to represent first-person sequential actions. Furthermore our finding is in agreement with the concurrent activation approach to sequential action (Estes, 1972) but not with integrative and chaining theories. It seems essential to develop a more comprehensive theory of (the development of) sequential action representation, which would need to address how novel components are integrated into a sequential plan, how the sequencing is generated, and whether this requires hierarchical representations. We are confident that the paradigm presented in chapter 8 can be helpful in answering some of these questions, especially by introducing further modifications of the task.

Due to the setup of our particular paradigm in chapter 3, we could not differentiate whether the action-effect knowledge that enabled goal perception came from prior first-person experience or whether prior observance of others' action effects was sufficient. Sommerville's et al. (2008) findings seem to suggest that self-experience is primary to observed action-effect perception. However, taken together with the findings of Sommerville et al. (2005) who used a conceptually close paradigm to ours, and a significant number of studies that used other methods (a review, see: Hauf, 2007; Elsner 2007; Melzoff, 2007), our finding suggests an important role for action-effect knowledge in third-person action perception. Furthermore, Paulus et al. (2012, 2013) showed that action-effects of observed actions elicit similar motor activity in the brain of 9-month-old infants as those of those of self-produced actions (in 8-month-old infants). This finding suggests that infants represent first-person- and third-person- action-effect knowledge in the same way. Thus these results are also compatible with theories, such as the Theory of Event Coding (Hommel et al., 2001), embodied cognition as a whole, and simulation theories (e.g., Meltzoff, 2007) that propose sensorimotor processes play an important role in the perception of third-person actions.

If one takes into account the functional and representational equivalence of self-performed and perceived actions as suggested by Theory of Event Coding (Hommel et al., 2001), this pattern fits with data suggesting that at 6 months of age infants can understand goal directed action (e.g. Woodward, 1998), or more accurately, experience violation-of-expectation to a change of goal, but are unable to perform true intentional action (distinguishing means from ends) until around 8 to 9 months of age (Goubet et al., 2006; Hauf, 2007; Piaget, 1936). Our findings also fit with results from studies on action perception, showing that infants at 9- but not 7-months of age can use observed action-effect relations to guide behavior (Hauf & Aschersleben, 2008). Additionally, our data suggest that motor resonance when listening to previously self-produced sounds in 8-months-olds, as found by Paulus et al. (2012), might indeed reflect the existence of knowledge about action-effect relations; and yet, we do not necessarily expect this knowledge to result in overt behavior, at least not at 7-months of age. Similar evidence for action-knowledge activation during action observation has been obtained in infants as young as 6 months (Nyström, 2008). Some authors have argued that it is lacking representational equivalence between self-produced actions and observed actions that prohibits infants younger than nine months from imitation (Hauf, 2007). Our data, together with those of Paulus et al. (2012, 2013), Verschoor et al. (2010), Nyström (2008), and Sommerville et al. (2005), suggest that it is not representational equivalence that is reached by 9 months of age, but the ability to successfully use bidirectional action-effect associations, learned either by observation or experience, for voluntary action.

Indeed, action-effect associations could even be a basis for simulation itself. One of the main problems of simulation theories is the problem of how “supra modal codes”, as Meltzoff (2007) calls them, can emerge. Supra modal codes refer to action representations that unite first- and third-person into a common representational framework. Such representations allow infants to see the behaviors of others as commensurate with their own. Based on a ‘like me’ perception of others, infants could use these codes to interpret third-person behavior (Meltzoff, 2007). Meltzoff suggest that this supra modal encoding space is with us from birth based on findings showing early imitation of facial gestures in newborn infants (Meltzoff & Moore, 1977, 1997). However other theorists would predict the emergence of this supramodal space thru experience (e.g., Hommel, 2003). Hommel (2003) hypothesizes that action-effects themselves bridge the gap between first- and third person action representation (as chapter 3 suggests). This would work something like this: if an infant performs a certain action this produces perceivable effects, if another person produces similar perceivable effects these effect trigger the infant's motor representation of their own action. This simple mechanism thus generates a certain perceptual equivalence in action representation of first- and third- person action (e.g., Paulus et al. 2012- 2013). Furthermore, unlike theories that propose that perception is translated automatically into supramodal codes, this theory can account for simulation of non-human actions.

Lastly, a finding from chapter 2 requires further discussion, the finding that by 7-months of age infants will *only* attribute a goal to an action if it is efficient, regardless to whether outcome selection information could disambiguate the action. In the discussion of chapter 2 the early presence of the preferential bias for means selection information is taken to suggest that the principle of rational action is a "core property" of the cognitive mechanisms for goal attribution. Although, Csibra and Gergely (1998) state that the principle of rational action is the initial state of the infants' naïve psychological theory, this "core property" does not necessarily have to be innate or come online via a process of maturation. It *could* also be learned. If one takes the perspective of embodied cognition and TEC seriously, one could go on to theorize in the following way. Infants, who have limited amounts of energy, are actors themselves. Therefore they should quickly learn, by evaluating action-effect contingencies, that acting efficiently toward a desired end state saves time and effort. If one further assumes that cognition emerges from sensory-motor processing, like embodied cognition does, one could assume that infants first internalize the principle of rational action themselves, and then generalize this principle to other agents.

Chapter 9

Final Conclusion

By using innovative paradigms, the present thesis provides convincing evidence that action-effect learning, and sensorimotor processes in general play a crucial role in the development of action-perception and production in infancy. This finding was further generalized to sequential action. Furthermore the thesis suggests that means-selection-, ends-selection information, and action-effect knowledge together feed into a unitary concept of goal. Both these findings have the potential to generate interesting new research question

References

Andrews, K. (2012). Mindreading animals: the debate over what animals know about other minds. [Review of the book: Mindreading animals: the debate over what animals know about other minds.] *Notre Dame philosophical reviews, an electronic journal.* Retrieved from http://ndpr.nd.edu/news/29824-mindreading-animals-the-debate-over-what-animals-know-about-other-minds.

Anisfeld, M. (1991). Review: neonatal imitation. *Developmental Review. 11*, 60–97.

Arcediano, F., Escobar, M., & Miller, R.R. (2005). Bidirectional associations in humans and rats. *Journal of Experimental Psychology: Animal Behavior Processes, 31*, 301-318.

Aschersleben, G., Hofer, T., & Jovanovic, B. (2008). The link between infant attention to goal directed action and later theory of mind abilities. *Developmental Science, 11*, 862-868.

Baillargeon R, Graber M, DeVos J & Black J. (1990). Why do young infants fail to search for hidden objects? *Cognition*, 36, 225–284.

Band, G.P.H., Van Steenbergen, H., Ridderinkhof, K.R., Falkenstein, M. & Hommel, B. (2009). Action-effect negativity: Irrelevant action effects are monitored like relevant feedback. *Biological Psychology, 82*, 211-218.

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

References

Baillargeon, R., Graber, M., DeVos, J., & Black, J. (1990). Why do young infants fail to search for hidden objects? *Cognition*, *36*, 255-284.

Baldwin, D.A., Baird, J.A., Saylor, M.M. & Clark, M.A., (2001). Infants parse dynamic action, *Child Development,* 72, 708–717.

Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition, 106*, 1382-1407.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unified theory of ADHD. *Psychological Bulletin, 121*, 65–94.

Barr, R., & Hayne, H. (1996). The effect of event structure on imitation in infancy: Practice makes perfect? *Infant Behavior and Development, 19*, 255-259.

Barr, R., Dowden, A., & Hayne, H. (1996). Developmental changes in deferred imitation by 6- to 24-month-old infants. *Infant Behavior & Development, 19*, 159–170.

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617-645.

Bauer, P. J., Hertsgaard, L. A., & Wewerka, S. S. (1995). Effects of experience and reminding on long-term recall in infancy: Remembering not to forget. *Journal of Experimental Child Psychology, 59,* 260–298.

Bauer, P. J., Hertsgaard L. A., Dropik P., Daly B. (1998). When even arbitrary order becomes important: developments in reliable temporal sequencing of arbitrarily ordered events. *Memory, 6*, 165–98.

References

Bauer, P. J., Wiebe, S. A., Waters, J. M., & Bangston, S. K. (2001). Reexposure breeds recall: effects of experience on 9-month-olds' ordered recall. *Journal of Experimental Child Psychology, 80*, 174–200.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91*, 276–292.

Beatty, J. & Lucero-Wagoner, B. (2000). The pupillary system. In J. Caccioppo, L.G. Tassinary, & G. Berntson (Eds.), *The handbook of psychophysiology* (2nd ed.) (pp. 142-162). Cambridge University Press.

Beisert, M., Zmyj, N., Liepelt, R., Jung, F., Prinz, W. & Daum, M. M. (2012). Rethinking 'rational imitation' in 14-month-old infants: a perceptual distraction approach. *PLoS ONE 7*(3): e32563. doi:10.1371/journal.pone.0032563.

Bellagamba, F., & Tomasello, M. (1999). Re-enacting intended acts: Comparing 12- and 18-month-olds. *Infant Behavior and Development*, *22*, 277–282.

Belsky, J., & Most, R. K. (1980). Free exploration to play: A cross-sectional study of infant free play behavior. *Developmental Psychology*, *17*, 630–639.

Bertenthal, B. I. (1996). Origins and early development of perception, action, and representation. *Annual Review of Psychology*, *47*, 431-459.

Bertenthal, B., & Clifton, R. K. (1998). Perception and action. In D. K. W. Damon & R. Siegler (Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception and language* (pp. 51–102). New York: Wiley.

References

Biro, S., Csibra, G., and Gergely, G. (2007). The role of behavioural cues in understanding goal-directed actions in infancy. In C. von Hofsten and K. Rosander (Eds.). *Progress in Brain Research:* From Action to Cognition. 164, 303-322.

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

Biro, S. & Vershoor, S. (2007). The role of efficiency of action in infants' goal attribution. Poster presented at the JPS conference, Amsterdam, NL.

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

Biro, S. Verschoor, S., Coalter, E. & Leslie, A. (under review). Outcome potential influences twelve-month-olds' interpretation of a novel action as goal-directed. *Cognitive Development.*

Blakemore, S.-J., Frith, C.D., & Wolpert, D.W. (1999). Spatiotemporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience*, *11*, 551–559.

Botvinick, M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review, 111*, 395–429.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology, 45*, 602-607.

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition, 83*, 223–240.

van den Broek, P. (1997). Discovering the cement of the universe: The development of event comprehension from childhood to adulthood. In P. van den Broek, P. Bauer, & T. Bourg (Eds.), *Developmental spans in event comprehension and representation: Bridging fictional and actual events* (pp. 321-342). Hillsdale, NJ: Erlbaum.

Butterworth, G., & Hopkins B. (1988). Hand–mouth coordination in the new–born baby. *British Journal of Developmental Psychology, 6*, 303–314.

Byrne, R.W., & Russon, A.E. (1998). Learning by imitation: a hierarchical approach. *Behavioral and Brain Sciences, 21*, 667-721.

Call, J. & M. Tomasello, 2008, Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Science*, 12: 187–92.

Cannon, E.N. & Woodward A.L (2010). Familiar actions trump action-effect" in goal detection in the first year. Paper presented at XVII. International Conference on Infant Studies (ICIS), March Baltimore.

Carey, S., & Spelke, E.S. (1994). Domain specific knowledge and conceptual change. In L.Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 169-200). New York: Cambridge University Press.

Cardoso-Leite, P., Mamassian, P., Schütz-Bosbach, S., & Waszak, F. (2010). A new look at sensory attenuation: Action-effect anticipation affects sensitivity, not response bias. *Psychological Science, 21*(12), 1740-1745.

Carmody, D.P., Dunn, S.M., Boddie-Willis, A.S., DeMarco, J.K., & Lewis, M. (2004). A quantitative measure of myelination development in infants, using MR images. *Neuroradiology, 46*, 781-786.

References

Carver, L. J. & Bauer P. J. (1999). When the event is more than the sum of its parts: 9-month-olds' long-term ordered recall. *Memory, 7*, 147-174.

Carver, L. J., & Bauer, P. J. (2001). The dawning of a past: the emergence of long-term explicit memory in infancy. *Journal of Experimental Psychology, 130*, 726–745.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* MIT Press.

Cocker, K. D., Fielder, A. R., Moseley, M. J., & Edwards, A. D. (2005). Measurements of pupillary responses to light in term and preterm infants. *Neuro-Ophthalmology, 29*, 95–101.

Cohen, L. B. (1972). Attention-getting and attention-holding processes of infant visual preferences. *Child Development, 43*, 869-879.

Corbetta, D., & Thelen, E (1999). Lateral biases and fluctuations in infants' spontaneous arm movements and reaching. *Developmental Psychobiology*, *34* (4), 237-255.

Csibra, G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science*, *1*, 255-259.

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

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

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

References

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

Claxton, L.J., Keen, R., & McCarty, M. E. (2003). Evidence of motor planning in infant reaching behavior. *Psychological Science, 14*, 354-356.

Clearfield, M. W., & Thelen, E. (2001). Stability and flexibility in the acquisition of skilled movement. In C. A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 253–266). Cambridge, MA: MIT Press.

Cleeremans, A., & McClellend, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General, 120*, 235–253.

Colombo, J., & Mitchell, D. W. (2009). Infant visual habituation. *Neurobiology of Learning and Memory, 92*, 225-234.

Constantinidis, C., Williams, G. V. & Goldman-Rakic, P. S. (2002). A role for inhibition in shaping the temporal flow of information in prefrontal cortex. *Nature Neuroscience, 5*, 175–180.

Conway C. M. & Christiansen M. H. (2001). Sequential learning in nonhuman primates. *Trends in Cognitive Science*, *5*, 539–546.

Cooper, R. P. & Shallice, T. (2006). Hierarchical schemas and goals in the control of sequential behavior. *Psychological Review, 113*, 887–916.

Corbetta, D., & Thelen, E (1999). Lateral biases and fluctuations in infants' spontaneous arm movements and reaching. *Developmental Psychobiology, 34* (4), 237-255.

Cox, R. F. A., & Smitsman, A. W. (2006). Action planning in young children's tool use. *Developmental Science, 9*, 628–641.

References

Davidson, D. (1985). Rational Animals. In E. Lepore & B. McLaughlin (Eds.) *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. New York: Basil Blackwell.

DeCasper, A. J. & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 1208*, 1174-1176.

Dell, G.S., Burger, L.K., & Svec, W.R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review, 104*, 123-147.

Dennett, D. (1987). *The Intentional Stance.* Cambridge, MA: MIT Press.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). London: Oxford University Press.

Diamond, A. (2006). The early development of executive functions. In E. Bialystok & F. Craik (eds.), *Lifespan Cognition: Mechanisms of Change* (pp. 70-95). NY: Oxford University Press.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 381*, 1387–1388.

Duncan, J. (1995). Attention, intelligence and the frontal lobes. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 721-733). Cambridge, MA: MIT Press.

Dutzi, I. B., & Hommel, B. (2009). The microgenesis of action-effect binding. *Psychological Research, 73*, 425-435.

Eenshuistra, R.M., Weidema, M.A., & Hommel, B. (2004). Development of the acquisition and control of action-effect associations. *Acta Psychologica, 115*, 185–209.

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 229–240.

Elsner, B., Hommel, B., Mentschel, C., Drzezga, A., Prinz, W., Conrad, B., & Siebner, H.R. (2002). Linking actions and their perceivable consequences in the human brain. *NeuroImage, 17*, 364-372.

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*, 732-751.

Elsner, B., & Hommel, B. (2004). Contiguity and contingency in the acquisition of action effects. *Psychological Research, 68*, 138-154.

Elsner, B. (2007). Infants' imitation of goal-directed actions: The role of movements and action effects. *Acta Psychol.* 124, 44–59.

Elsner, B., Hauf, P., & Aschersleben, G. (2007). Imitating step by step: a detailed analysis of 9- to 15-month olds' reproduction of a 3-step action sequence. *Infant Behavior & Development*, *30*, 325-335.

Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, DC: Winston.

Del Giudice, M., Manera, V., & Keysers, C. (2009). Programmed to learn? The ontogeny of mirror neurons. *Developmental Science, 12,* 350-363.

Fabbri-Destro, M. & Rizzolatti, G. (2008). Mirror neurons and mirror systems in monkeys and humans. *Physiology, 23*, 171-9.

Fantz, R. (1958). Pattern vision in young infants. *The Psychological Record, 8*, 43–47.

References

Fenker, D., Waldmann, M.R., & Holyoak, K. J. (2005). Accessing causal relations in semantic memory. *Memory & Cognition, 33,* 1036-1046.

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience, 9,* 878-879.

Falck-Ytter, T. (2008). Face inversion effects in autism: A combined looking time and pupillometric study. *Autism Research, 1*, 297-306.

Fawcett, C. & Gredebäck, G. (2013). Infants use social context to bind actions together into a collaborative sequence *Developmental Science, 16*, 841-849.

Fodor, J. A. (1975). *The Language of Thought*, Cambridge, Massachusetts: Harvard University Press.

Fodor, J.A. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind.* Cambridge, MA: MIT Press.

Foster, S. M., Kisley, M. A., Davis, H. P., Diede, N. T., Campbell, A. M., & Davalos, D. B. (2013). Cognitive function predicts neural activity associated with pre-attentive temporal processing. *Neuropsychologia, 51*, 211-219.

Friedman, S., Nagy, A. N., & Carpenter, G. C. (1970). Newborn attention: Differential response decrement to visual stimuli. *Journal of Experimental Child Psychology, 10,* 44–51.

Fuster, J. (2002). Physiology of executive functions: The perception–action cycle. In: Stuss, D. T. & Knight, R. T. (Eds.). *Principles of frontal lobe function.* (pp. 96–108). New York: Oxford University Press.

Fuster, J. M. (1993). Frontal lobes. *Current Opinion in Neurobiology, 3,* 160–165.

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*, 1239–1255.

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

Gergely, G., & Watson, J. S. (1999). Early social-emotional development: Contingency perception and the socialbiofeedback model. In P. Rochat (Ed.), *Early social cognition* (pp. 101–136). Hillsdale: Erlbaum Publishers.

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

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

Gergely, G. (2009). Kind of Agents: The origins of understanding instrumental and Communicative Agency. U. Goshwami, (Ed.). *Blackwell Handbook of Childhood Cognitive Development. 2nd Edition.* Oxford: Blackwell Publishers.

Gervain, J., Berent, I., & Werker, J.F. (2012). Binding at birth: the newborn brain detects identity relations and sequential position in speech. *Journal of Cognitive Neuroscience, 24(3)*, 564-574.

Gibson JJ. (1979). *The ecological approach to visual perception.* Boston: Houghton-Mifflin.

Gopnik, A. & Meltzoff, A. N. (1997). *Words, Thoughts and Theories*, Cambridge, MA: MIT Press.

References

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., & Kushnir, T. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111*, 3-23.

Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics, 56*, 544–549.

Goubet, N., Rochat, P., Maire-Leblond, C. & Poss, S. (2006). Learning from others in 9- to 18-month-old infants. *Infant Child Development, 15*, 161-177.

Gredebäck, G., Stasiewicz, D., Falck-Ytter, T., Rosander, K., & von Hofsten, C. (2009) Action Type and Goal Type Modulate Goal-Directed Gaze Shifts in 14-Month-Old Infants. *Developmental Psychology, 4*, 1190-1194.

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition, 114*, 197-206.

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review, 77*, 73-99.

Guajardo, J. J. & Woodward, A. L. (2004). Is Agency Skin Deep? Surface Attributes Influence Infants' Sensitivity to Goal-Directed Action. *Infancy*, 6(3), 361-384.

Hanuschkin, A., Ganguli, S. & Hahnloser, R. H. R. (2013). A Hebbian learning rule gives rise to mirror neurons and links them to control theoretic inverse models, *Frontiers in Neural Circuits, 7*, 106.

Harless, E. (1861). Der Apparat des Willens. *Zeitschrift fuer Philosophie und philosophische Kritik, 38*, 50–73.

References

Hartshorn, K., Rovee-Collier, C., Gerhardstein, P., Bhatt, R.S., Wondoloski, T.L., Klein, P., Gilch, J. Wurzel, N., & Campos-de-Carvalho, M. (1998). The ontogeny of long-term memory over the first year-and-a-half of life. *Developmental Psychobiology, 32,* 69-89

Hauf, P., Elsner, B. & Aschersleben, G. (2004). The role of action effects in infants' action control. *Psychological Research, 68,* 115–125.

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

Hauf, P. (2007). Infant's perception and production of intentional actions. In C. von Hofsten & K. Rosander (Eds.), Progress in Brain Research: *From Action to Cognition, 164*, 285-301.

Hauf, P. & Aschersleben, G. (2008). Action-effect anticipation in infant action control. *Psychological Research, 72,* 203-210.

Hausman, D. (1998). *Causal asymmetries*. Cambridge: Cambridge University Press.

Hayek, F.A. (1952). *The Sensory Order*. London: Routledge & Kegan Paul.

Heider, F, & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57,* 243–259.

Heineman-Pieper, J., & Woodward, A. (2003). Understanding infants' understanding of intentions: two problems of interpretation. *Consciousness and Cognition*, 12, 770–772.

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

Henderson, A. M. E., & Woodward, A. L. (2011). ‘Let’s work together’: what do infants understand about collaborative goals? *Cognition, 121* , 12–21.

Henry, F. M., & Rogers, D. E. (1960). Increased response latency for complicated movements and a ‘‘memory drum’’ theory of neuromotor reaction. *Research Quarterly, 31,* 448–458.

Henson, R. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology, 36*, 73-137.

Herwig, A., & Horstmann, G. (2011). Action-effect associations revealed by eye movements. *Psychonomic Bulletin & Review, 18,* 531-537.

Hernik M. & Southgate V (2012). Nine-months-old infants do not need to know what the agent prefers in order to reason about its goals: On the role of preference and persistence in infants’ goal-attribution. *Developmental Science, 15*, 714–722.

Hess, E. H., & Polt, J. (1960). Pupil size as related to interest value of visual stimuli, *Science, 132*, 149–150.

Hess, E. H., & Polt, J. (1964). Pupil size in relation to mental activity during simple problem-Solving. *Science, 13*, 1190-1192.

Hess, E. H. (1975). *The tell-tale eye: how your eyes reveal hidden thoughts and emotions*. New York: Van Nostrand Reinhold.

Heyes, C. M. (2001). Causes and consequences of imitation. *Trends in Cognitive Sciences, 5*, 253-261.

Heyes C. M. (2010). Where do mirror neurons come from? *Neuroscience and Biobehavioural Reviews, 34*, 575–583.

References

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience, 21,* 1229–1243

Hofer, T., Hauf, P. & Aschersleben, G. (2005). Infant's perception of goal-directed actions performed by a mechanical device *Infant Behavior & Development, 28,* 466–480.

Hommel, B. (1996). The cognitive representation of action: automatic integration of perceived action effects. *Psychological Research, 59*, 176-186.

Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding: a framework for perception and action planning. *Behavioral and Brain Sciences, 24*, 849–937.

Hommel, B. (2003). Planning and representing intentional action. *The scientific world journal, 3*, 593-608.

Hommel, B. (2004). Coloring an action: Intending to produce color events eliminates the Stroop effect. *Psychological Research, 68*, 74-90.

Hommel, B. (2005). Perception in action: multiple roles of sensory information in action control. *Cognitive Processing*, *6*, 3-14.

Hommel, B., & Elsner, B. (2009). Acquisition, representation, and control of action. In E. Morsella, J. A. Bargh, & P. M. Gollwitzer (eds.), *The psychology of action,* pp. 371-398 New York: Oxford University Press.

Houghton, G. & Tipper, S.P. (1996). Inhibitory mechanisms of neural and cognitive control: applications to selective attention and sequential action. *Brain and Cognition, 30*, 20-43.

Huestegge, L., & Kreutzfeldt, M. (2012). Action effects in saccade control. *Psychonomic Bulletin and Review, 19*, 198-203.

Hughes, G., Desantis, A., & Waszak, F. (2013a). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction and motor prediction. *Psychological Bulletin, 139*, 133-151.

Hughes, G., Desantis, A., & Waszak, F. (2013b). Attenuation of auditory N1 results from identity-specific action-effect prediction. *European Journal of Neuroscience. 37*, 1152-1158.

Hull, C. L. (1931). Goal attraction and directing ideas conceived as habit phenomena. *Psychological Review, 38*, 487-506.

Hume, D. (1964). *Treatise of human nature* (edited by L. A. Selby-Bigge). London: Oxford University Press. (Original work published 1739).

Hupe, J. M., Lamirel, C., & Lorenceau, J. (2009). Pupil dynamics during bistable motion perception. *Journal of Vision 9*, 1–19.

Jacob, P. (2012). Sharing and ascribing goals. *Mind and Language, 27*, 202-229.

Jackson, I., & Sirois, S. (2009). Infant cognition: Going full factorial with pupil dilation. *Developmental Science, 12*, 670-679.

James, W. (1890). *The principles of psychology*. New York: Dover Publications.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40*, 1–19.

Johnson, S. C. (2000). The recognition of mentalistic agents in infancy. *Trends in Cognitive Sciences,* Volume 4, Issue 1, 22-28.

References

Johnson S.C., Shimizu Y. A., & Ok S. (2007). Actors and actions: The role of agent behavior in infants' attribution of goals. *Cognitive Development*, 22, 310–322.

Johnson, S. C., Ok, S.-J., & Luo, Y. (2007). The attribution of attention: 9-month-olds' interpretation of gaze as goal-directed action. *Developmental Science, 10*, 530-537.

Jones, S. S. (2009). The development of imitation in infancy. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 2325–2335.

Jovanovic, B., Kiraly, I., Elsner, B., Gergely, G., Prinz, W., & Aschersleben, G. (2007). The role of effects for infants'perception of action goals. *Psychologia, 50*, 273-290.

Kahneman, D. & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154*, 1583-1585.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, New York: Prentice-Hall.

Kayed, N. S., & Van der Meer, A. L. H. (2000). Timing strategies used in defensive blinking to optical collisions in 5- to 7-month-old infants. *Infant Behavior and Development*, *23*, 253–270.

Keen, R. (2005) Using perceptual representations to guide reaching and looking. Chapter in John Rieser, Jeffrey Lockman, & Charles. Nelson (Eds.) *Action as an organizer of learning and development: Minnesota Symposia on Child Psychology, 33* (pp. 301-322). Mahwah, NJ: Lawrence Erlbaum Associates.

Kenward, B. (2010). 10-month-olds visually anticipate an outcome contingent on their own action. *Infancy*, *15*, 337–361.

Kiesel, A., & Hoffmann, J. (2004). Variable action effects: Response control by context-specific effect anticipations. *Psychological Research., 68*, 155-162.

Kim, R., Seitz, A. R., Feenstra, H., Shams, L. (2009). Testing assumptions of statistical learning: Is it long-term and implicit? *Neuroscience Letters, 461*,145-149.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*, B35-B42.

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

Kiraly, I., Jovanovic, B., Prinz, W., Aschersleben, G., & Gergely, G. (2003). The early origins of goal attribution in infancy. *Consciousness and Cognition, 12*, 752-769.

Klein, P.J., & Meltzoff, A.N. (1999). Long-term memory, forgetting, and de-ferred imitation in 12-month-old infants. *Developmental Science, 2*, 102–113.

Klein, A., Hauf, P. & Aschersleben, G. (2006). The role of action effects in 12-month olds' action control: A comparison of televised model and live model. *Infant Behavior and Development, 29*, 535-544.

de Kleijn R, Kachergis G and Hommel B (2014) Everyday robotic action: lessons from human action control. *Frontiers in Neurorobotics, 8*, 13.

Koch, I., Keller, P. E., & Prinz, W. (2004). The ideomotor approach to action control: Implications for skilled performance. *International Journal of Sport and Exercise Psychology, 2,* 362-375.

Kray, J., Eenshuistra, R., Kerstner, H., Weidema, M., & Hommel, B. (2006). Language and action control: The acquisition of action goals in early childhood. *Psychological Science*, *17*, 737-741.

Kühn, S., Keizer, A., Rombouts, S.A.R.B., & Hommel, B. (2011). The functional and neural mechanism of action preparation: Roles of EBA and FFA in voluntary action control. *Journal of Cognitive Neuroscience, 23*, 214-220.

Kunde, W. (2003). Temporal response–effect compatibility. *Psychological Research*, *67*, 153–159.

Laeng, B., & Falkenberg, L. (2007). Women's pupillary responses to sexually significant others during the hormonal cycle. *Hormones and Behavior, 52,* 520–530.

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

Leslie, A.M. (1982). The perception of causality in infants. *Perception, 11* (2), 173–186.

Leslie, A.M. (1984). Infant perception of a manual pick-up event. *British Journal of Developmental Psychology*, *2*, 19–32.

Leslie, A. M. (1994). ToMM, ToBy, and agency: Core architecture and domain specificity, In: Hirschfeld, L.A., Gelman, S.A. (Eds.), *Mapping the mind: Domain specificity in cognition and culture*. Cambridge University Press, pp. 119-148.

Leslie, A. M. (2010). Agency, Action and Choice. Paper presented at the Opening Conference of Cognitive Developmental Center at Central European University, Budapest, Hungary. http://www.youtube.com/watch?v=Km3es5aNvoY

Lewkowicz, D. J. (2004a). Serial order processing in human infants and the role of multisensory redundancy. *Cognitive Processing, 5*, 113–122.

References

Lewkowicz, D. J. (2008). Perception of dynamic & static audiovisual sequences in 3- and 4-month-old infants. *Child Development, 79*, 1538-1554.

Lewkowicz, D. J., & Berent, I. (2009). Sequence learning in 4 month-old infants: Do infants represent ordinal information? *Child Development, 80*, 1811-1823.

Li, K. Z. H., Lindenberger, U., Rünger, D**.**, & Frensch, P. A. (2000). The role of inhibition in the regulation of sequential action. *Psychological Science, 11(4)*, 343-347.

Libby, W. L., Lacy, B. C., & Lacy, J. I. (1973). Pupillary and cardiac activity during visual attention. *Psychophysiology, 10*, 270-294.

Lotze, R. H. (1852). *Medicinische Psychologie oder die Physiologie der Seele.* Leipzig: Weidmann'sche Buchhandlung.

Lukowski, A. F., Wiebe, S. A., Haight, J. C., Waters, J. M., Nelson, C. A., & Bauer, P. J. (2005). Forming a stable memory representation in the first year of life: Why imitation is more than child's play. Developmental Science, 8, 279-298.

Luo, Y. & Baillargeon, R. (2005). Can a Self-propelled Box Have a Goal? Psychological Reasoning in 5-Month-Old Infants. *Psychological Science,* 16(8), 601-608.

Luo, Y. & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition,* 105, 489–512.

Luo, Y. & Johnson C.S. (2009). Recognizing the role of perception in action at 6 months. *Developmental Science,* 12(1), 142–149.

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

References

Luria, A. R. (1973). The frontal lobes and the regulation of behavior. In K. H. Pribram (Ed.), *Psychophysiology of the frontal lobes* (pp. 3–26). Oxford, UK: Academic Press.

Lurz, R. W. (2009). Animal minds. Internet encyclopedia of philosophy. 2012. From http://www.iep.utm.edu/ani-mind/#SH1d.

Lurz, R. W. (2011). *Mindreading Animals: The Debate over What Animals Know about Other Minds*. Cambridge MA, MIT Press.

Mandler, J. M. & McDonough, L. (1995). Long-term recall of event sequences in infancy. *Journal of Experimental Child Psychology, 59*, 457-474.

Mäntylä, T., Carelli, M. G., & Forman, H.(2007). Time monitoring and executive functioning in children and adults. *Journal of Experimental Child Psychology*, *96*, 1-19.

Marcus, G. F., Fernandes, K. J., & Johnson, S. P. (2007). Infant rule learning facilitated by speech. *Psychological Science, 18*, 387-391.

Marcovitch, S., & Lewkowicz, D. J. (2009). Sequence learning in infancy: The independent contributions of conditional probability and pair frequency information. *Developmental Science, 12*, 6, 1020-1025.

McCarty, M. E., Clifton, R. K., & Collard, R. R. (1999). Problem solving in infancy: The emergence of an action plan. *Developmental Psychology, 35*, 1091–1101.

McCarty, M. E., Clifton, R. K., Ashmead, D. H., Lee, P., & Goubet, N. (2001) How infants use vision for grasping objects. *Child Development, 72*, 973-987.

McCormack, T., & Atance, C. M. (2011). Planning in young children: A review and synthesis. *Developmental Review, 31*, 1-31.

Melcher, T., Weidema, M., Eenhuistra, R. M., Hommel, B., & Gruber, O. (2008). The neural substrate of the ideomotor principle: An event-related fMRI analysis. *Neuroimage, 39*, 1274–1288.

Meltzoff, A.M. & Moore, M.K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198*, 75-78.

Meltzoff, A. N. (1988). Infant Imitation After a 1-Week Delay: Long-Term Memory for Novel Acts and Multiple Stimuli, *Developmental Psychology., 24,* 470–476.

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

Meltzoff, A. N., & Moore, M. K. (1997). Explaining facial imitation: a theoretical model. *Early Development and Parenting, 6*, 179–192.

Meltzoff, A.N. (2002). Imitation as a mechanism of social cognition: origins of empathy, theory of mind, and representation of action. In *Blackwell's Handbook of Childhood Cognitive Development* (Goswami, U., ed.), pp. 6–25, Blackwell.

Meltzoff, A. N., & Prinz, W. (Eds.) (2002). *The imitative mind: Development, evolution and brain bases.* Cambridge, UK: Cambridge.

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

Michel, G. F. (1998). A lateral bias in the neuropsychological functioning of human infants. *Developmental Neuropsychology, 14* (4), 445-469.

Miltner, W. H. R., Braun, C. H. & Coles, M. G. H. (1997). Event-related potentials following incorrect feedback in a time-estimation task: evidence for a ‘generic’ neural system for error detection. *Journal of Cognitive Neuroscience, 9,* 788–798.

Montare, A. (1977). Human temporal behavior and discrimination reversal learning. *Pavlovian Integrative Physiological and Behavioral Science, 12*, 232–246.

Moore, M.K., & Meltzoff, A.N. (2004). Object permanence after a 24-hr delay and leaving the locale of disappearance: the role of memory, space, and identity. *Developmental Psychology, 40*, 606–620.

Müsseler, J. & Prinz, W. (2002) (Hrsg.). *Lehrbuch Allgemeine Psychologie.* Heidelberg: Spektrum Akademischer Verlag.

Namy, L.L., Campbell, A.L., & Tomasello, M. (2004). The changing role of iconicity in non-verbal symbol learning: A U-shaped trajectory in the acquisition of arbitrary gestures. *Journal of Cognition and Development*, *5*, 37-57.

Needham, A., Barrett, T. & Peterman, K. (2002).A pick-me-up for infants’ exploratory skills: Early simulated experiences reaching for objects using ‘sticky mittens’ enhances young infants’object exploration skills. *Infant Behavior & Development, 25*, 279–295.

Norman, D. A. & Shallice, T. (1986). Attention to action: willed and automatic control of behaviour. In R. J. Davidson, G.E. Schwartz and D. Shapiro (Eds.) *Consciousness and Self Regulation: Advances in Research*, Vol. IV pp1-18, New York: Plenum.

Nyström, P. . (2008). The infant mirror neuron system studied with high density EEG. *Social Neuroscience, 3*(3), 334.

Osaka, N., Ikeda, T., Osaka, M. (2012). Effect of Intentional Bias on Agency Attribution of Animated Motion: An Event-Related fMRI Study. *PLoS ONE, 7*(11): e49053. doi:10.1371/journal.pone.0049053.

Olofson, E. L., & Baldwin D. (2011). Infants recognize similar goals across dissimilar actions involving object manipulation. *Cognition, 118,* 258-264.

Paulus, M., Hunnius, S., & Bekkering, H. (2011). Can 14- to 20-month-old children learn that a tool serves multiple purposes? A developmental study on children's action goal prediction. *Vision Research, 51*, 955-960.

Paulus, M. (2011). *Development of action perception: Neurocognitive mechanisms underlying children's processing of others' actions.* Nijmegen: Ipskamp.

Paulus, M., Hunnius, S., Vissers, M., and 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*, 901–910.

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

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

Paulus, M., Hunnius, S., van Elk, M., & Bekkering, H. (2012). How learning to shake a rattle affects 8-month-old infants’ perception of the rattle’s sound: Electrophysiological evidence for action-effect binding in infancy. *Developmental Cognitive Neuroscience, 2*, 90-96.

Paulus, M., Hunnius, S., & Bekkering, H. (2013). Neurocognitive mechanisms subserving social learning in infancy: Infants' neural processing of the effects of others' actions. *Social Cognitive and Affective Neuroscience, 8,* 774-779.

Paulus, M. (2014). How and why do infants imitate? An ideomotor approach to social and imitative learning in infancy. *Psychonomic Bulletin & Review,* published online 28th of February 2014.

Piaget, J. (1936, 1963). *The origins of intelligence in children*. New York: W.W. Norton & Company, Inc.

Piaget, J. (1954). *The construction of reality in the child.* New York: Basic Books,

Piaget, J. (1962). *Play, dreams and imitation in childhood.* New York: W. W. Norton & Co.

Picozzi, M., de Hevia, M. D., Girelli, L., & Macchi-Cassia, V. (2010). Seven-month-old infants detect ordinal numerical relationships within temporal sequences. *Journal of Experimental Child Psychology, 107,* 359-367.

Posner, M. I., & Cohen, Y. A. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance* (pp. 531–554). Hillsdale, NJ: Erlbaum

Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral Brain Science, 4,* 515-526.

Prinz, W. (1990). A common coding approach to perception and action. In Neumann O., Prinz W. (Eds.), *Relationships between perception and action* (pp. 167-201). Berlin: Springer.

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

References

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research, 91*, 176-180.

Ravenscroft, I. (1997). Folk Psychology as a Theory. *The Stanford Encyclopedia of Philosophy* (Fall 2010 Edition), Edward N. Zalta (ed.), Retrieved from http://plato.stanford.edu/archives/fall2010/entries/folkpsych-theory.

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

Rizzolatti G, Craighero L (2004). The mirror neuron system. *Annual Review Neuroscience*, *27*, 169-192.

Rochat, P., & Striano, T. (1999). Emerging self-exploration by 2 month-old infants. *Developmental Science, 2,* 206–218.

Rochat, P. (2001). *The infant's world.* Cambridge: Harvard University Press.

Romberg, A. R. & Saffran, J. R. (2013). Expectancy learning from probabilistic input by Infants. *Frontiers in Psychology, 3*, 610.

Rosenbaum, D. A. (1987). Successive approximations to a model of human motor programming. *Psychology of Learning and Motivation, 21*, 153–182.

Rosenbaum, D.A., Marchak, F., Barnes, H.J., Vaughan, J., Slotta, J., & Jorgensen, M. (1990). Constraints for action selection: overhand versus underhand grips. In: Jeannerod, M. (ed.), Attention and performance XIII (pp. 321–342). Hillsdale, NJ: Erlbaum.

Rovee-Collier, C. (1987). “Learning and memory in infancy,” in *Handbook of Infant Development*, ed. J. D. Osofsky (New York: Wiley), 98–148. 2nd Edn.

Rovee-Collier, C. (1996). Shifting the focus from what to why? *Infant Behavior and Development, 19,* 385–400.

Rovee-Collier, C. (1999). The development of infant memory. *Current Directions in Psychological Science, 8,* 80-85.

Rovee, C. K., & Rovee, D. T. (1969). Conjugate reinforcement of infant exploratory behavior. *Journal of Experimental Child Psychology, 8,* 33–39.

Rumelhart, D. E.,& Norman, D. 1982. Simulating a skilled typist: A study of skilled cognitivemotor performance. *Cognitive Science, 6*, 1–36.

Sakai, K., Kitaguchi, K., & Hikosaka, O. (2003). Chunking during human visuomotor sequence learning. *Experimental Brain Research, 152*, 229–242.

Sato, A. & Itakura, S. (2013). Intersubjective action-effect binding: Eye contact modulates acquisition of bidirectional association between our and others' actions. *Cognition, 127,* 383-390.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user’s guide.* Pittsburgh: Psychology Software Tools Inc.

Scholl, B. & Leslie, A. (1999). Modularity, Development and ‘Theory of Mind’, *Mind and Language*, *14,* 131–53.

Schrama, T. and Biro, S. (2005). Baby-timer: LabVIEW-based program for looking time studies. University of Leiden.

Scerif, G., Karmiloff-Smith, A., Campos, R., Elsabbagh, M., Driver, J., & Cornish, K. (2005). To look or not to look? Typical and atypical development of oculomotor control. *Journal of Cognitive Neuroscience, 4,* 591-604.

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

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

Shimpi, P., 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*, 309-329.

Shin, Y. K., Proctor, R. W., & Capaldi, E .J. (2010). A review of contemporary ideomotor theory. *Psychological Bulletin, 136*, 943-974.

Snodgrass, G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for naming agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174-215.

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

Sodian, B. (2011). Theory of mind in infancy. *Child Development Perspectives, 5,* 39-43.

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

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

Sommerville, J. A., Woodward, A. L. & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition, 96*, B1-B11.

Sommerville, J.A., Hildebrand, E., & Crane, C.C. (2008). Experience matters: The impact of doing versus watching on infants' subsequent perception of tool use events. *Developmental Psychology*, *44*, 1249–1256.

Sommerville, J.A. & Crane, C.C. (2009). Ten-month-old infants use prior information to identify an actor's goal. *Developmental Science, 12,* 314–325.

Sternberg, S., Monsell, S, Knoll, R.L., & Wright, C.E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typing. In G. E. Stelmach (Ed.) *Information Processing in Motor Control and Learning.* New York: Academic Press. Pp. 117-152.

Stock, A., & Stock, C. (2004). A short history of ideo-motor action. *Psychological Research, 68*, 176-188.

Stöcker, C. & Hoffmann, J. (2004). The ideomotor principle and motor sequence acquisition: tone effects facilitate movement chunking. *Psychological Research, 68,* 126–137.

Suanda, S., Tompson, W., & Brannon, E. M. (2008). Changes in the ability to detect ordinal numerical relationships between 9 and 11 months of age. *Infancy, 13*, 308–337.

Teinonen, T., Fellmann, V., Näätänen, R., Alku, P., Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience, 10*, 21.

Tissaw, M.A. (2007). Making sense of neonatal imitation. *Theory & Psychology. 17*, 217-242.

Tolman, E.C., & Brunswik, E. (1935). The organism and the causal texture of the environment. *Psychological Review, 42*, 43-77.

Tomasello, M., Davis-Dasilva, M., Camak, L. & Bard, K. (1987). Observational learning of tool use by young chimpanzees. *Human evolution, 2*, 175-183.

Tomasello, M. (1999). *The cultural origins of human cognition.* Cambridge: Harvard University Press.

Van Hof, P., Van der Kamp, J., & Savelsbergh, G. J. P. (2006). Three- to eight-month-old infants' catching under monocular and binocular vision. *Human Movement Science, 25*, 18–36.

Verschoor, S. A., Weidema, M., Biro, S., & Hommel, B. (2010). Where do action goals come from? Evidence for spontaneous action-effect binding in infants. *Frontiers in Cognition, 1*, 201.

Verschoor, S. & Biro, S. (2012). Means selection information overrides outcome selection information in infants' goal attribution. *Cognitive Science, 36*, 714–725.

Verschoor, S. A., Eenshuistra, R. M., Kray, J., Biro, S. & Hommel, B. (2012). Explicit learning of arbitrary and non-arbitrary action–effect relations in adults and 4-year-olds. *Frontiers in Psychology, 2*, 354.

Verschoor, S., Spapé, M., Biro, S., Hommel, B. (2013). From outcome prediction to action selection: Developmental change in the role of action-effect bindings. *Developmental Science, 16,* 801-814.

Verschoor, S. A., Eenshuistra, R.M., Weidema M.A., & Hommel, B. (submitted). *When words (and thoughts) speak louder than actions: Explicit task representation suppresses implicit action-effect associations.* Manuscript submitted for publication.

von Hofsten, C. (2004). An action perspective on motor development. *Trends in Cognitive Sciences, 8*, 266-272.

Vygotsky, L. S. (1985). *Thought and Language.* Cambridge, MA: The M.I.T. Press.

Wagner, L. & Carey, S. (2005). 12-month-old infants represent probable endings of motion events. *Infancy*, *7*(1), 73–83.

Waldmann, M.R. (1996). Knowledge-based causal induction. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.). *The psychology of learning and motivation*: *Vol. 34. Causal learning* (pp. 47-88). San Diego: Academic Press.

Waldmann, M. R., & Martignon, L. (1998). A Bayesian network model of causal learning. In M. A. Gernsbacher & S. J. Deny (Eds.), *Proceedings of the 20th Annual Conference of the Cognitive Science Society* (pp. 1102-1107). Mahwah, NJ: Erlbaum.

Wang, Q., Bolhuis, J., Rothkopf, C.A, Kolling, T., Knopf, M., & Triesch, J. (2012) Infants in Control: Rapid Anticipation of Action Outcomes in a Gaze-Contingent Paradigm. *PLoS ONE*, *7*(2).

Watson, J. S. (1967). Memory and 'contingency analysis' in infant learning. *Merrill-Palmer Quarterly, 13*, 55-76.

Wessel, J. R., Danielmeier, C. & Ullsperger, M. (2011) Error Awareness Revisited: Accumulation of Multimodal Evidence from Central and Autonomic Nervous Systems. *Journal of Cognitive Neuroscience*, *23*, 3021–3036.

Wellman, H. M., Phillips, A. T., Dunphy-Lelii, S., & Lalonde, N. (2004). Infant social attention predicts preschool social cognition. *Developmental Science, 7*, 283–288.

References

Wellman, H. M., Lopez-Duran, S., LaBounty, J., & Hamilton, B. (2008). Infant attention of intentional action predicts preschool social cognition. *Developmental Psychology*, *44*, 618–623.

Wentworth, N., Haith, M. M., & Hood, R. (2002). Spatiotemporal regularity and interevent contingencies as information for infants' visual expectations. *Infancy, 3*, 303 – 321.

Willatts, P. (1999). Development of means-end behavior in young infants: pulling a support to retrieve a distant object. *Developmental Psychology, 35*, 651–666.

Wolfe, C. D., & Bell, M. A. (2007). The integration of cognition and emotion during infancy and early childhood: Regulatory processes associated with the development of working memory. *Brain and Cognition, 65*, 3–13.

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

Woodward, A. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development*, 22(2), 145-160.

Woodward, A. & Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science*, 11(1), 73-77.

Woodward, A.L., Sommerville, J.A.& Guajardo, J.J. (2001) How infants make sense of intentional action. In Malle, B.; Moses, L.; Baldwin, D., editors. *Intentionality: A key to human understanding.* (pp146-169) Cambridge MA, MIT Press.

Woodward, A. L., Sommerville, J. A., Gerson, S., Henderson, A. M. E. & Buresh, J. (2009). The emergence of intention attribution in infancy. *Psychology of Learning and Motivation, 51*, 187-222.

References

Woodward, A. (2009). Infants’ Grasp of Others’ Intentions. *Current Directions in Psychological Science*, *18*, 53-57.

Zélanti, P.& Droit-Volet, S. (2011). Cognitive abilities explaining age-related changes in time perception of short and long durations. *Journal of Experimental Child Psychology, 109*, 143–157.

Zelazo, P. R., Brody, L. B., & Chaika, H. (1984). Neonatal habituation and dishabituation of headturning to rattle sounds. *Infant Behavior and Development*, *7,* 311-321.

Ziessler, M. (1994). The impact of motor responses on serial pattern learning. *Psychological Research, 57*, 30-41.

Ziessler, M. (1998). Response-effect learning as a major component of implicit serial learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24(4)*, 962–978.

Samenvatting

In dit proefschrift worden de resultaten van het onderzoek “Learning about goals: Development of action perception and action control” beschreven. De mens is gedwongen zich aan te passen aan zijn immer veranderende omstandigheden. Dit doet de mens door doelen na te streven. Gezien de constante veranderingen die onze omgeving ondergaat en de flexibiliteit waarmee de mens hierop kan reageren, is het zeer onwaarschijnlijk dat zulke doelen aangeboren en onveranderlijk zijn. Aangezien de mens zonder het vermogen om doelen te stellen niet zou kunnen overleven, is de vraag hoe de mens doelen stelt en intentioneel handelt allerminst triviaal.

Aangenomen dat doelen niet (allemaal) zijn aangeboren, rijst de vraag hoe de mens nieuwe doelgerichte handelingen kan leren. In principe is het antwoord op deze vraag op het eerste gezicht simpel: de mens leert doelgericht te handelen door te leren van wat anderen doen en door eigen ervaring. Echter als men wat kritischer naar dit antwoord kijkt komen nieuwe vragen op zoals “Hoe leren mensen van anderen?”. Om te leren van de doelgerichte handelingen van anderen moet de mens doelgerichte handeling en toevallige beweging kunnen onderscheiden. “Hoe neemt de mens een handeling van een ander als doelgericht waar?” is dus een logische vervolgvraag. Aangezien communicatie onmogelijk is zonder waarneming van andermans doelen, en de mens van nature sociaal is, is het vermogen om doelen te onder?”.

Aangezien het vermogen doelgericht te handelen zich ontwikkelt gedurende de eerste paar levensjaren is het een logische stap om doelgericht handelen te onderzoeken bij zuigelingen. In deze dissertatie worden achtereenvolgens de volgende vragen onderzocht “ Hoe leren zuigelingen doelgerichte handelingen van anderen waar te nemen?” en “Hoe leren zuigelingen zelf doelgerichte handelingen uit te voeren?” door verschillende theorieën te toetsen.

Doelgericht handelen leren door anderen te observeren

Deze theorieën stellen verschillende soorten informatie voor als criterium voor zuigelingen om te beoordelen of een handeling doelgericht is. De zogenaamde “Teleological stance” theorie stelt dat de manier waarop een handeling wordt uitgevoerd ten opzichte van de omgeving een cruciale factor is om te bepalen of het gaat om intentioneel handelen. Meer associatief georiënteerde theorieën voorspellen dat het herhaaldelijk dezelfde uitkomst kiezen uit verschillende mogelijkheden een cruciale factor is. In de hoofdstukken 2, 3 en 4 word onderzocht welke van deze twee soorten informatie belangrijker is voor het waarnemen van doelen op basis van geobserveerde handelingen, de manier waarop een handeling wordt uitgevoerd of welke uitkomst er wordt gekozen. De resultaten van deze onderzoeken suggereren dat de manier waarop de handeling wordt uitgevoerd (in dit geval dat de handeling wordt aangepast aan de omstandigheden) belangrijker is dan welke keuze er wordt gemaakt. Bovendien suggereren de resultaten dat zowel keuze van uitkomst, als manier van uitvoering *dezelfde* doelattributie beïnvloeden. Dit laat zien dat zuigelingen al vroeg in plaats van twee losstaande concurrerende concepten, een abstract concept van doelgerichtheid hebben waarin beide vormen van informatie een rol spelen.

In hoofdstuk 5 worden twee andere theorieën tegen elkaar getoetst. De motor resonantie theorie beweert dat het cruciaal is om ervaring te hebben met een bepaalde handeling om deze bij anderen als doelgericht te kunnen waarnemen. Anderzijds benadrukt de Ideomotor theorie dat handelingen alleen als doelgericht worden beoordeeld als ze opvallende effecten teweeg brengen. Hoofdstuk 4 laat zien dat zuigelingen die zelf ervaring opdeden met een voor hen nieuwe handeling die opvallende effecten veroorzaakte, een dergelijke geobserveerde handeling (zonder effecten) als doelgericht waarnemen.

Terwijl zuigelingen die weliswaar dezelfde ervaring kregen maar dan zonder opvallend effect, dezelfde handeling niet als doelgericht waarnemen. Dit onderzoek toont aan dat ervaring met een bepaalde handeling op zich niet voldoende is om deze handeling, wanneer deze wordt uitgevoerd door iemand anders, te zien als doelgericht. Naast de ervaring met de handeling is het cruciaal om te hebben ervaren dat een dergelijke handeling opvallende effecten kan veroorzaken.

Nieuwe doelen leren door ervaring

Het tweede deel van de dissertatie behandelt het aanleren van nieuwe doelen door ervaring. De Ideomotor theorie behelst dat men eerst informatie moet vergaren over samenhangen tussen bewegingen en hun gevolgen om deze later intentioneel te kunnen toepassen. Hoe dit werkt illustreer ik aan de hand van een voorbeeld. Als een baby voor het eerst een rammelaar ziet weet hij/zij niet dat deze kan rammelen. De baby zal uit gezonde interesse de rammelaar gaan onderzoeken. De kans is groot dat op een gegeven moment de baby per toeval een beweging maakt (actie) die een rammel geluid tot gevolg heeft (effect). De Ideomotor theorie beweert dat actie en effect vanaf nu aan elkaar verbonden zijn via een verbinding die twee kanten opgaat. Dit betekent dat wanner de baby weer dezelfde beweging (met de rammelaar) maakt, hij/zij het rammel geluid verwacht. Andersom betekent het ook dat als de baby aan het rammelgeluid denkt hij/zij automatisch ook denkt aan de bijbehorende (rammel) actie. Door dit principe kunnen zuigelingen door middel van exploratie een arsenaal aan actie-effect verbindingen opdoen die later gebruikt kunnen worden om doelgerichte handelingen uit te voeren.

Er is veel evidentie gevonden voor het ideomotor principe in volwassenen en er werd vaak zonder direct bewijs aangenomen dat het ook toepasbaar is op kinderen. De doelstelling van de volgende vier hoofdstukken is te bewijzen dat de Ideomotor theorie toepasbaar is op de ontwikkeling van doelgericht handelen. In hoofdstuk 5 wordt aangetoond dat actie-effect leren op een vergelijkbare manier werkt in 4 jarigen en volwassenen. In hoofdstuk 6 wordt voor het eerst bewijs geleverd dat het Ideomotor principe ook toepasbaar is op zuigelingen van 12 maanden, maar nog niet op 9 maanden oude zuigelingen. Aangezien we hadden verwacht dat ook 9 maanden oude zuigelingen actie-effect leren zouden laten zien (omdat deze wel al doelgericht kunnen handelen) zijn we in hoofdstuk 7 op zoek gegaan naar een betere methode om actie-effect leren te meten bij zuigelingen. In dit hoofdstuk hebben we actie-effect leren onderzocht door middel van oogbewegingen en pupillometrie. Met deze methode waren we wel in staat om actie-effect leren aan te tonen bij 9 maanden oude zuigelingen. Bovendien konden we met deze methode een directe vergelijking maken tussen volwassenen en zuigelingen. In hoofdstuk 8 hebben we deze nieuwe methode verder ontwikkeld om hem ook toepasbaar te maken voor het vrijwel totaal onontgonnen onderzoek naar het aanleren van sequentiële acties.

Samenvattend laat dit proefschrift zien dat zuigelingen al vroeg een abstract concept van doelgerichtheid ontwikkelen waarvoor ze gebruik maken van verschillende soorten informatie. Verder is de Ideomotor theorie gevalideerd als een theorie die grote verklarende waarde heeft voor de ontwikkeling van doelgericht handelen.

Curriculum Vitae

Stephan Verschoor was born in Alphen aan den Rijn on the 16th of June 1975. He attended secondary education in Waddinxveen. From 1995 to 2006, he studied Psychology at Leiden University. After completing his Psychology thesis on action-effect learning under supervision of Rena Eenshuistra, he graduated in August 2006. He worked as a research assistant for Rena Eenshuistra and Prof. Dr. Bernhard Hommel at the chair of sociology cognitive psychology, from 2004 until 2006.

From 2006 to 2012, he did his PhD research which was supervised by Dr. Szilvia Biro and Prof. Dr. Bernhard Hommel (Cognitive Psychology, Leiden University & Leiden Institute for Brain and Cognition). This research focused on the development of action perception and action control. He used a variety of behavioral and neuroscientific techniques to study the neurocognitive mechanisms that sub serve action control and action perception. The results of his PhD research are described in this thesis.

Stephan Verschoor currently works as an assistant professor at the LMU in Munich. Here, he cooperates with Prof. Dr. Beate Sodian and Prof. Dr. Markus Paulus and, among other things, investigates the development of sequential action control and its cognitive substrate across the lifespan.

Curriculum Vitae

Dankwoord

Graag wil ik allereerst mijn promotor Bernhard en co-promotor Szilvia hartelijk bedanken voor hun steun, opbouwende kritiek en geduld. Eveline, Niels en Markus wil ik bedanken voor het lezen van dit proefschrift. De leden van de oppositiecommissie bedank ik hier ook alvast. Daarnaast wil ik natuurlijk mijn onmisbare coauteurs bedanken.

Verder wil ik mijn collega's bedanken voor gezelligheid, steun, advies en ideeen. In willekeurige volgorde noem ik hier: Michiel, Anne, Rena, Maaike, Henk, Ludger, Esther, Rinus, André, Roy, Nelleke, Pascal, Jesse, Marieke, Mikael, Marc, Stephen, Gezinus, Guido, Wido, Gert, Sander, Fenna, Niels, Kees, William, Soghra, Kalinka, Jun, Eveline en Lorenza. Albertien, Atie, Eva en Marianne, jullie mentale en secretariële ondersteuning was uitstekend! Ook wordt de technische ondersteuning door Thijs, Henk, en Michiel zeer gewaardeerd. Bovendien was dit alles onmogelijk geweest zonder de medewerking van proefleiders en de vaak schattige proefpersoontjes.

Verder ben ik vooral Nicolien, mijn familie, vrienden en vriendinnen uiterst dankbaar voor de onmisbare steun en liefde!