Learning about goals : development of action perception and action control

Verschoor, S.A.

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

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)

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.

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).

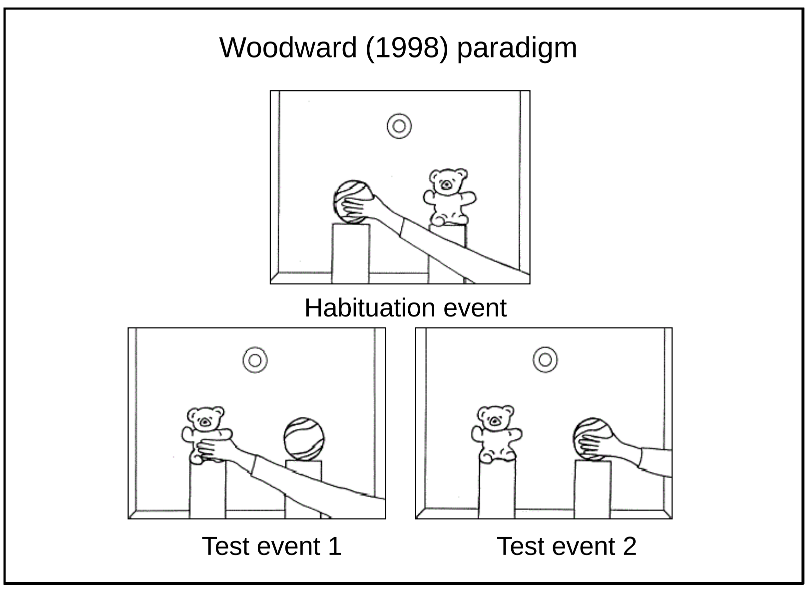

Figure 1. Woodwards 1998 paradigm. Infants were habituated to a hand grasping a toy at one of two locations. From habituation to test the toys change sides. Infants dishabituate more to a change of goal (Test event 1) than to a change of location (Test event 2) indicating the infants expects the actor 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).

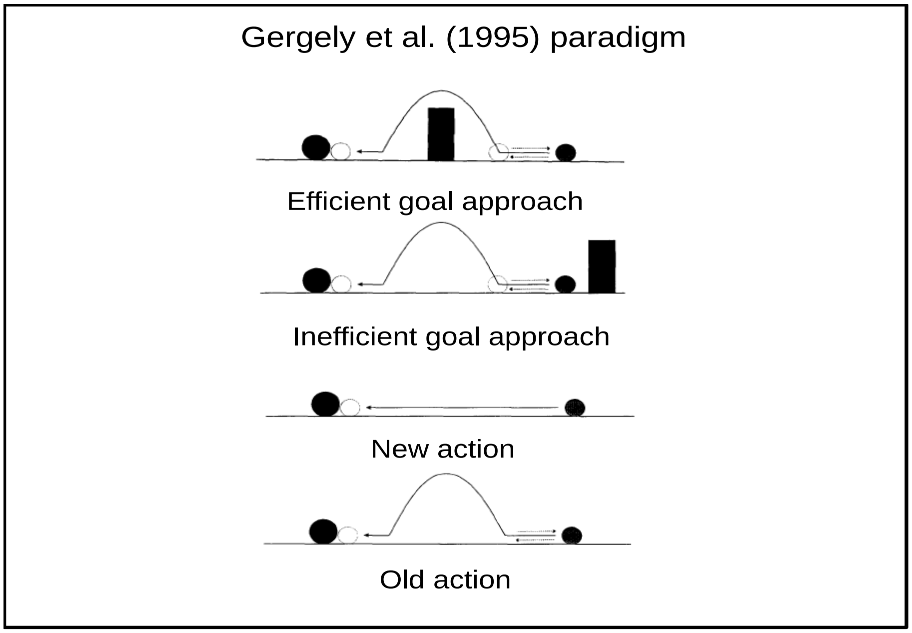

Figure 2. Gergely et al. (1995) paradigm. Infants were habituated to either the efficient or the inefficient goal approach wherein a small circle approaches a big circle (in the inefficient condition the obstacle was not in between the two circles yet the circle still jumped). During test they were shown 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.

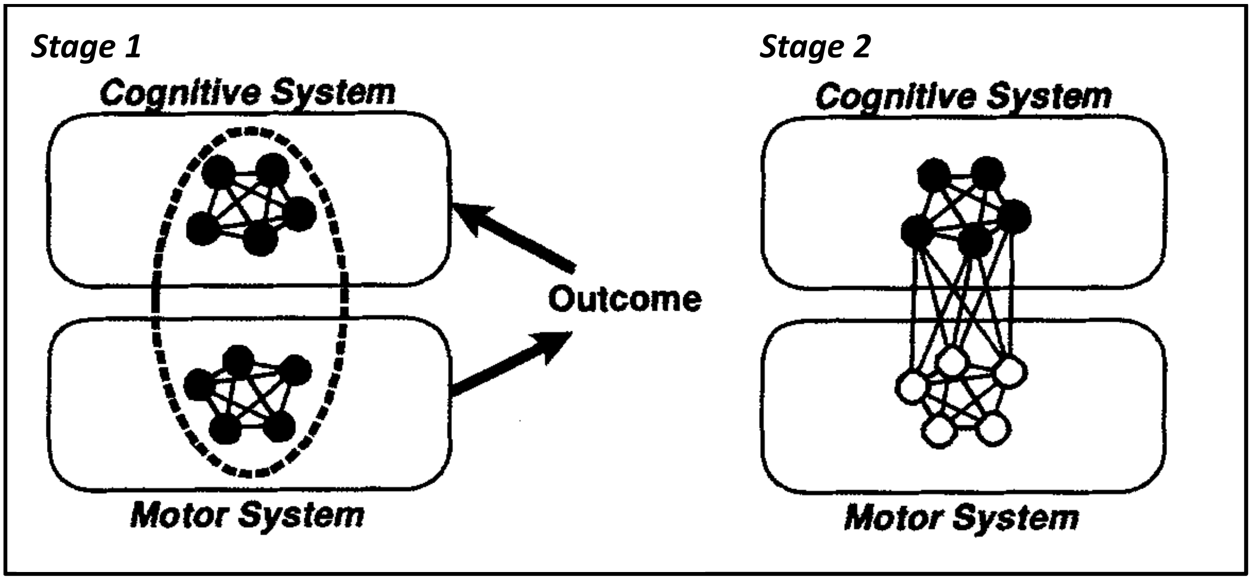

Figure 3. Two stage model by Elner and Hommel (2001). At stage1, the motor patter producing a particular effect is automatically integrated with the cognitive codes representing this effect. At stage 2, the motor pattern is intentionally executed by activating the cognitive codes that represent its effect.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

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.