



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



S.A. Verschoor^{a,b,*}, M. Paulus^a, M. Spapé^c, S. Biro^{b,d}, B. Hommel^b

^a Developmental Psychology Unit, LMU München, Germany

^b Leiden University Institute for Psychological Research & Leiden Institute for Brain and Cognition, Netherlands

^c Helsinki Institute for Information Technology, Finland

^d Center for Child and Family Studies, Leiden University, Netherlands

ARTICLE INFO

Article history:

Received 13 June 2013

Revised 15 January 2015

Accepted 20 January 2015

Keywords:

Ideomotor action

Action control

Sequential action

Cognitive development

Eye tracking

Pupillometry

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 oculo-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.

© 2015 Elsevier B.V. All rights reserved.

1. 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 & Sommerville, 2000; Woodward, Sommerville, Gerson, Henderson, & Buresh, 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.

1.1. 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

* Corresponding author at: Leopoldstrasse 13, 80802, München, Germany.

E-mail address: VerschoorSA@gmail.com (S.A. Verschoor).

and the actions producing them (Hommel, 1996; Hommel, Müssele, Aschersleben, & Prinz, 2001; Prinz, 1990, 1997). 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, Weidema, & Hommel, 2004; Kray, Eenshuistra, Kerstner, Weidema, & Hommel, 2006) and adults, suggesting the mechanism responsible to be fast-acting (Dutzi & Hommel, 2009), automatic (Band, van Steenberghe, Ridderinkhof, Falkenstein, & Hommel, 2009; Elsner & Hommel, 2001), 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).

Until recently, research on the importance of action effects for infants mainly focused on third-person action interpretation (e.g., Biro & Leslie, 2007; Hauf, 2007; Kiraly, Jovanovic, Prinz, Aschersleben, & Gergely, 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, Hauf, & Aschersleben, 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; Meltzoff, 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, Hunnius, van Elk, & Bekkering, 2012). The youngest infants showing evidence for reversing bidirectional action effects for action control are 9-month-olds (Verschoor, Weidema, Biro, & Hommel, 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.

1.2. 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, Haith, & 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).

1.3. Sequential action representation in infancy

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 (Baldwin, Baird, Saylor, & Clark, 2001; Conway & Christiansen, 2001; Lewkowicz, 2004). VOE studies report that around the age of 6 months infants start to evaluate the efficiency of sequential actions (Biro, Verschoor, & Coenen, 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 (2011) 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, Wang, Matz, & Woodward, 2013;

Henderson & Woodward, 2011) 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). Adding 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 and 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 et al., 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 the rudimentary abilities to learn from and interpret third-person sequential action, as well as the abilities to plan and control first-person sequential action emerge by the end of the first year. The studies further suggest that temporal organization, action effects and goals are important sources of information, yet they leave open how such information is used for integrating action steps into coherent sequences. In the current study we will attempt to clarify the cognitive mechanism responsible for this feat.

1.4. 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 Fig. 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.

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 et al., 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

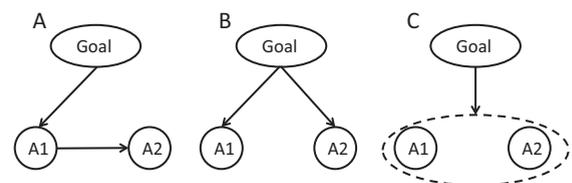


Fig. 1. Models of sequential action. Schematic representation of activation in A: Chaining models of sequential action, activation cascades forward through the different elementary actions, B: Concurrent models of sequential action, all elementary actions are activated simultaneously where after completion 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 (adapted from Müsseler & Prinz, 2002).

mechanisms to be too slow to account for the speed of practiced sequential action (e.g. [Lashley, 1951](#); [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 ([Houghton & Tipper, 1996](#); [Posner & Cohen, 1984](#)). The concurrent model can account for the empirical finding of more prospective than retrospective intrusion errors ([Dell, Burger, & Svec, 1997](#); [Li, Lindenberger, Rüniger, & 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 ([Cox & Smitsman's, 2006](#); [McCarty et al., 1999](#)).

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](#); [Sakai, Hikosaka, & Nakamura, 2004](#)). 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](#)).

1.5. 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 actions, with element R1 preceding

element R2. All models imply that priming or otherwise activating R1 might spread activation to R2, but they differ regarding their predictions when R2 is primed/activated. [James' \(1890\)](#) chaining model would not predict that priming R2 leads to activation of R1, since the sequence is assumed to be represented by unidirectional effect–response bindings ($R1 \rightarrow R2$). [Greenwald's \(1970\)](#) version would predict the spreading of activation from R2 to R1, as sequences are represented by associations formed between the effects of the elementary actions. Conversely, concurrent activation models would predict that activating R2 leads to the inhibition of R1, 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 R1. 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 et al. \(2013\)](#) 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 ([Falck-Ytter, Gredebäck, & von Hofsten, 2006](#); [Gredebäck & Melinder, 2010](#); [Senju & Csibra, 2008](#)). The paradigm records Reaction Times (RTs) and Response Frequencies (RFs). The study of [Verschoor et al. \(2013\)](#) and other recent studies that demonstrated saccade-effect learning in adults ([Herwig & Horstmann, 2011](#); [Huestegge & Kreutzfeldt, 2012](#)), 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, 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](#);

Laeng1, Sirois, & Gredebäck, 2012; Verschoor et al., 2013). TERPs indicate motivational phenomena such as increased 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). 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 R^a and R^b were followed by action effects E^a and E^b ($R^a \rightarrow E^a$; $R^b \rightarrow E^b$). Performing action R^a in response to (or as a result of being primed by) effect E^b in the test phase ($E^b \rightarrow R^a$) would be considered congruent, while performing the same action in response to E^a ($E^a \rightarrow R^a$) would be considered incongruent.

Introducing sequences that consist of two action components (components 1 and 2) renders the definition somewhat more complicated (see Table 1). Our participants were exposed to two sequences of actions (A and B) and their effects during acquisition: $R1^A \rightarrow E1^A \rightarrow R2^A \rightarrow E2^A$ and $R1^B \rightarrow E1^B \rightarrow R2^B \rightarrow E2^B$. In the test phase, we presented the action effect of one of the second action components ($E2^A$ or $E2^B$) and we tested whether this would affect processes related to the first action components ($R1^A$ and $R1^B$). If infants represent the experienced action sequences as a unity, cueing the effect of the second element ($E2^A$ or $E2^B$) could affect the activation of the first action elements ($R1^A$ or $R1^B$). The pairings of effect $E2^A$ and action $R1^A$ ($E2^A \rightarrow R1^A$) or of effect $E2^B$ and action $R1^B$ ($E2^B \rightarrow R1^B$) were considered congruent, and the pairings of $E2^A$ and action $R1^B$ ($E2^A \rightarrow R1^B$) or of $E2^B$ and action $R1^A$ ($E2^B \rightarrow R1^A$) incongruent.

Finding any difference depending on congruency would provide evidence for a coherent cognitive representation of sequential action in infants. Crucially, 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.

Table 1

The action sequences learned during acquisition and the congruent and incongruent responses during test.

Acquisition	Test	
	Congruent	Incongruent
Sequential action		
$R1^A \rightarrow E1^A \rightarrow R2^A \rightarrow E2^A$	$E2^A \rightarrow R1^A$	$E2^A \rightarrow R1^B$
And	And	And
$R1^B \rightarrow E1^B \rightarrow R2^B \rightarrow E2^B$	$E2^B \rightarrow R1^B$	$E2^B \rightarrow R1^A$

2. Methods

2.1. 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.

2.2. 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 cm (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 (1280×1024 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 optimal accuracy of $.5^\circ$ and allows for a certain amount of head movement by the subjects ($30 \times 22 \times 30$ cm). It recorded gaze direction and pupil-size. Stimulus presentation was controlled by a PC running E-prime® software (Schneider, Eschman, & Zuccolotto, 2002).

2.3. Stimuli

The visual stimuli used were as follows (see Fig. 2). The background color of the screen was gray. 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, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis, & Morton, 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 200 ms effect sounds which were equalized on loudness, "tring" and "piew" (Verschoor, Eenshuistra, Kray, Biro, & Hommel, 2012; Verschoor et al., 2013) and complex high- and low- note sound waves of 1574- and 776-Hz.

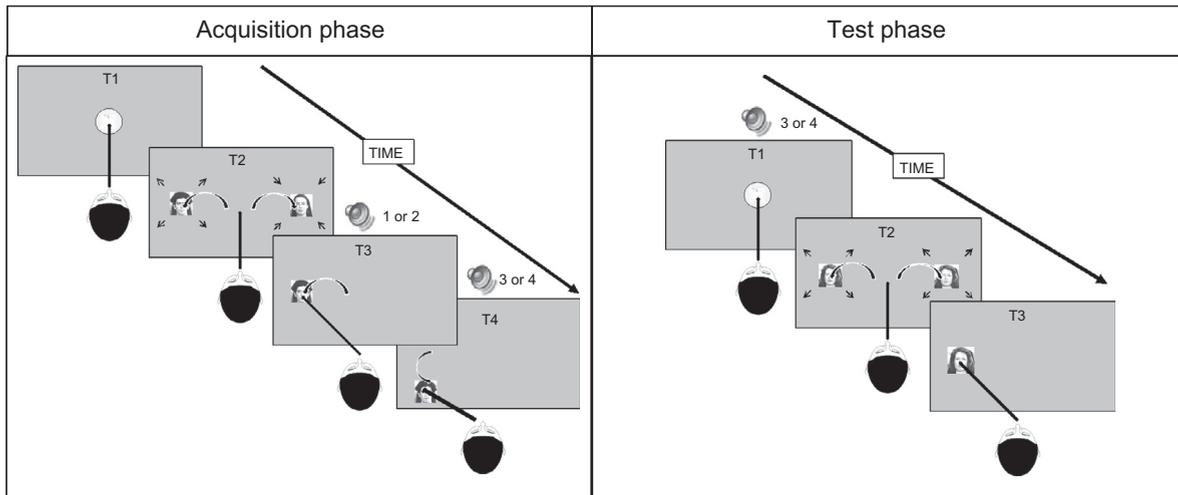


Fig. 2. *Acquisition trial:* Each trial starts with an intertrial interval of 500 ms. T1: A fixation dot is displayed at screen center. T2: After successful fixation, faces appear at either side of the screen where they started to pulsate. T3: Depending on the saccade target, the face at the other side disappears and an effect sound is played for 200 ms. T4: Depending on which side was chosen the face moves up or down whereafter a second saccade is made and a second sound effect is played for 200 ms. *Test trial:* Each trial starts with an intertrial interval of 500 ms. T1: A fixation dot is displayed at screen center. After successful fixation one of the previous action effects is played. T2: The dot disappears whereafter the same face appears on both sides. T3: The participant freely chooses where to saccade.

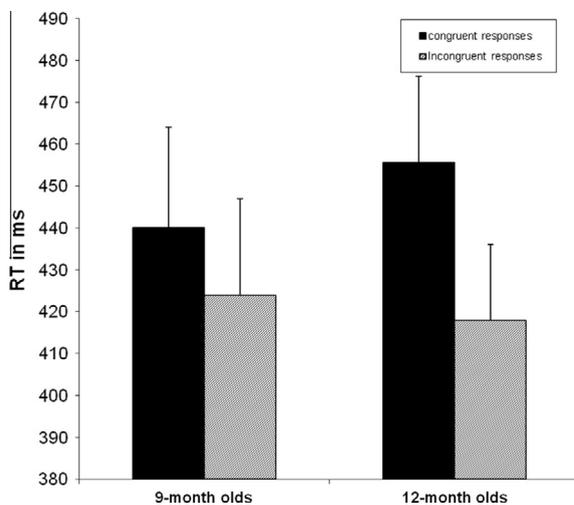


Fig. 3. Mean reaction times (+SE) for 9-month-olds ($N = 14$) and 12-month-olds ($N = 16$) in congruent and incongruent test trials.

2.4. 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.

2.4.1. Acquisition phase

The experiment began with an acquisition-phase of 36 trials (see Fig. 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 ($R1^A \rightarrow E1^A \rightarrow R2^A \rightarrow E2^A$ or $R1^B \rightarrow E1^B \rightarrow R2^B \rightarrow E2^B$). Each saccade sequence consisted of two distinct actions, first one to the left or right ($R1^A$ or $R1^B$) whereafter an up- or downward action followed ($R2^A$ or $R2^B$, 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 200 ms effect-sounds ("tring" or "pieuw") was presented

($E1^A$ or $E1^B$, 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 200 ms effect-sounds ($E2^A$ or $E2^B$, "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 2000 ms; if by then no response was detected, the trial was repeated. After each trial, an inter-trial-interval of 500 ms was used.

2.4.2. Test phase

The test phase of 32 trials followed directly afterwards (see Fig. 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 200 ms during which an effect-sound was presented that was previously triggered by one of the two secondary eye-movements ($E2^A$ or $E2^B$). 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 ($R1^A$ or $R2^B$). The question of interest was whether the direction of these saccades ($R1^A$ or $R2^B$) would be biased by the tones ($E2^A$ or $E2^B$). Except for absence of

auditory effects after the saccades, the remaining procedure was as during acquisition.

2.5. 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 2000 ms before the presentation of the sound onset or RT, to 8000 ms 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 500 ms after effect presentation (see Fig. 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 .03 mm in 17 ms. Gaze data were recorded in pixel coordinates, averaged between eyes and filtered using a 10 Hz 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

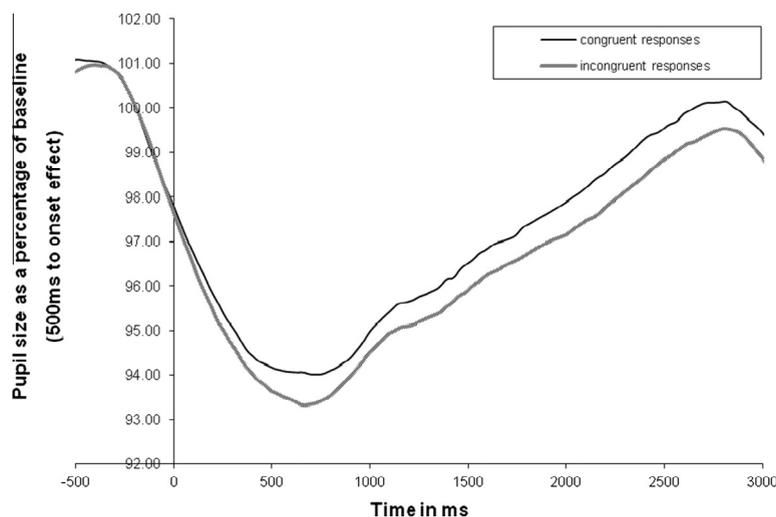


Fig. 4. Relative pupil sizes for congruent and incongruent responses to baseline, stimulus-locked.

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.

3. Results and discussion

3.1. 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 2). 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, $F(1,25) = 6.25$, $p = .019$, $\eta^2p = .20$. Separate analyses showed no main effect in 9-month-olds (RT-left = 999 ms, RT-right = 1046 ms) and a tendency toward faster right-ward responses in 12-month-olds, $F(1,12) = 3.84$, $p = .074$, $\eta^2p = .24$ (RT-right = 982 ms, RT-left = 1089 ms). Secondly, vertical response location interacted with age group, $F(1,25) = 4.63$, $p = .04$, $\eta^2p = .16$. Separate analyses showed no main effect in 9-month-olds (RT-up = 1008 ms, RT-down = 1037 ms) and a tendency toward faster downward responses in 12-month-olds, $F(1,12) = 3.82$, $p = .07$, $\eta^2p = .24$ (RT-up = 1089, RT-left = 982 ms). 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, $F(1,28) = 175$, $p < .001$, $\eta^2p = .86$, indicating responses toward contracting pictures were slower (partial RT-contracting = 603 ms, partial RT-expanding = 428 ms).

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., Kaye & 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.

3.2. 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 ($p > .4$).

3.2.1. 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.

3.2.2. 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 (see Table 1 for mapping details) as within-subjects factor, revealed a significant effect indicating 29 ms-slower responses for congruent trials, $F(1,28) = 4.15$, $p = .05$, $\eta^2p = .13$; the interaction with age was not significant ($p > .3$) (see Fig. 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 $F(1,15) = 5.47$, $p = .03$, $\eta^2p = .27$. 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 ($F(1,12) = 4.96$, $p = .05$, $\eta^2p = .29$).

Our main finding is: cueing of the secondary action of the action sequence interfered with executing its first, as evidenced by the longer RT's for congruent responses in the 12-month-olds (see Table 1 for mapping details).

Table 2

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)

Table 3

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

Age group 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)

Results were less clear in the 9-month-olds. The signed rank test did not show significant results while using Number Of Completed Acquisition Trials (NOCAT) as a covariate in the RT analysis resulted in a significant effect in 9-month-olds. This suggests that the NOCAT was an important factor for the strength of the effect in this age group whereas the 12-month-olds showed a ceiling effect for the NOCAT necessary for the uptake of the sequential action. 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.

3.2.3. 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 (see Table 1 for mapping details) responses TERPs used a 500 ms 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 0 ms), p 's > .7. Adults' TERPs start from 200 to 300 ms after stimulus onset and peak in the range of 500 ms to 2000 ms (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 500–2000 ms post effect onset. An ANOVA with MART as covariate revealed that, overall, participants exhibited larger relative pupil dilations during congruent responses, $F(1,27) = 4.12$, $p = .05$, $\eta^2p = .13$, 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 1000–2500 ms post effect onset time window. Again, pupil dilations were significantly larger in congruent trials, $F(1,25) = 5.03$, $p = .03$, $\eta^2p = .16$, independently of age, $p > .09$ (see Fig. 4).

For the response-locked analysis, we calculated the percentage of dilation from a 700-ms time window starting at saccade onset, to the same 500 ms pre-stimulus baseline. An ANOVA with MART as covariate yielded a tendency for larger relative dilation in congruent trials, $F(1,27) = 3.51$, $p = .07$, $\eta^2p = .12$, while the interaction with age group was not significant, $p > .8$ (see Fig. 5). Adding NOCAT as additional covariate resulted in a significant effect ($F(1,26) = 5.48$, $p = .03$, $\eta^2p = .17$), again without an interaction with age group ($p > .9$).

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.

3.2.4. Gaze position

In the test phase we primed the second action of the two-element sequences carried out in the acquisition phase by presenting the corresponding action effect (see

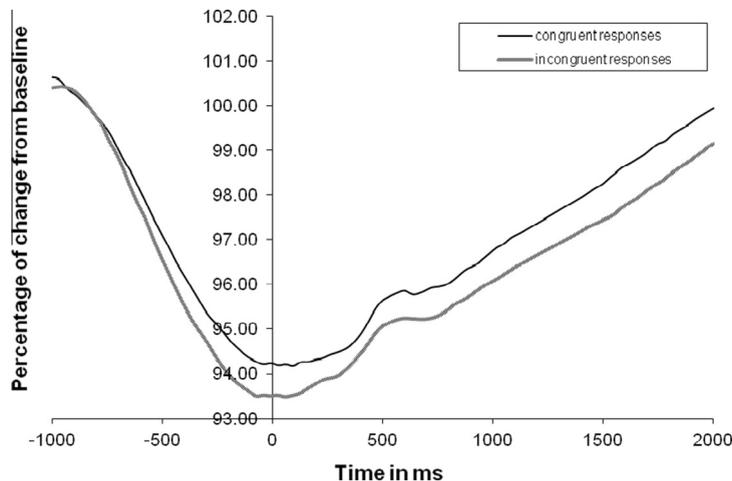


Fig. 5. Relative pupil sizes for congruent and incongruent responses to baseline, response locked.

Table 1 for mapping details). 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 650 ms thereafter (corresponding to the mean RT plus mean random ITI) and compared these segments to a 150 ms 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, $F(1,27) = 4.83$, $p = .04$, $\eta^2p = .15$ (effect size = 22 pixels; see Fig. 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 ($F(1,27) = 4.70$, $p = .04$, $\eta^2p = .15$) and did not vary with age ($p = .2$), but not significant for congruent responses ($p > .26$).

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.

4. 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 et al., 2001; Carver & Bauer, 1999, 2001; Elsner et al., 2007; Lukowski et al., 2005). 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).

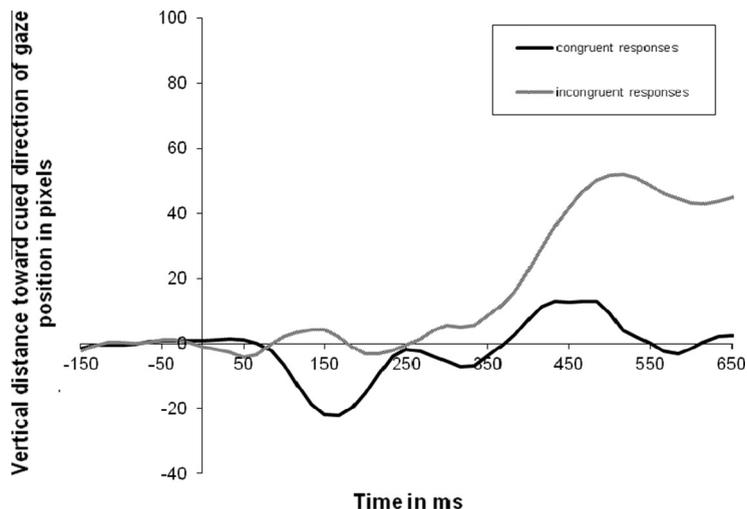


Fig. 6. Vertical distance from midline toward cued direction of gaze position in pixels for congruent and incongruent responses. Time is 0, is the moment the effect starts.

The fact that evaluation of third-person sequential action is apparent significantly earlier in development, in 6- to 7-month-olds (Baillargeon, Graber, DeVos, & Black, 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; Cox & Smitsman, 2006; McCarty et al., 1999) and cognitive theories (e.g. Botvinick & Plaut, 2004; Constantinidis 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, Barnett, Thomas, & Munro, 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 (Mäntylä, Carelli, & Forman, 2007; Zélandi & Droit-Volet, 2011). 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 Stocker and Hoffmann (2004) found that action effects facilitate chunking (for a similar point see Ziessler & Nattkemper, 2001). Furthermore, the activation of the secondary vertical component in the incongruent trials 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 (see for a similar point: Herbolt & Butz, 2012; Kachergis, Wyatte, O'Reilly, De Kleijn, & Hommel, 2014), 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 & McClelland, 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; Cox & Smitsman, 2006; McCarty et al., 1999) 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.

One could question whether our findings are generalizable to other action systems (manual, postural, etc.). Ideomotor theory makes no distinction between effectors and effect modalities, and ideomotor motivated research does not suggest such a distinction. Furthermore, saccade-effect learning is now established in adults (Herwig & Horstmann, 2011; Huestegge & Kreutzfeldt, 2012) and

infants (Verschoor et al., 2013) suggesting that the oculomotor system is controlled in the same way as for instance the system for manual action control. We acknowledge that saccadic eye movements operate on very short timescales where priming and inhibition may play a larger role than in other types of actions that involve gross movements of the body (e.g. reaching and locomotion) that operate on relatively slower timescales. Future research will have to clarify this issue.

Altogether it remains essential to develop 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 further modifications of our paradigm will help to increase insight into the general cognitive mechanisms underlying action planning (e.g., by cueing the first action component and examining how this affects the availability of the second) 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.

Acknowledgements

This research was supported by the Netherlands Organization for Scientific Research. We thank the reviewers for their useful suggestions and especially thank Thijs Schrama for technical support and Henk van Steenbergen for analytical support.

References

- Baillargeon, R., Graber, M., DeVos, J., & Black, J. (1990). Why do young infants fail to search for hidden objects? *Cognition*, 36, 255–284.
- Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, 113(3), 329–349.
- Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition*, 106, 1382–1407.
- Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development*, 72, 708–717.
- 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.
- 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 & Development*, 19, 255–259.
- 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, 165–198.
- 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., 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. Cacioppo, L. G. Tassinari, & G. Berntson (Eds.), *The handbook of psychophysiology* (2nd ed., pp. 142–162). Cambridge University Press.
- Biro, S., & Leslie, A. (2007). Infants' perception of goal-directed actions. Development through cues-based bootstrapping. *Developmental Science*, 10, 379–398.
- Biro, S., Verschoor, S., & Coenen, L. (2011). Evidence for a unitary goal-concept in 12 months old infants. *Developmental Science*, 14(6), 1255–1260.
- 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.
- 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.
- Claxton, L. J., Keen, R., & McCarty, M. E. (2003). Evidence of motor planning in infant reaching behavior. *Psychological Science*, 14, 354–356.
- Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General*, 120, 235–253.
- Cohen, L. B. (1972). Attention-getting and attention-holding processes of infant visual preferences. *Child Development*, 43, 869–879.
- 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.
- Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition*, 107, 705–717.
- de Kleijn, R., Kachergis, G., & Hommel, B. (2014). Everyday robotic action: Lessons from human action control. *Frontiers in Neurorobotics*, 8, 13.
- 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.
- Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, 381, 1387–1388.
- 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.
- 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. (2007). Infants' imitation of goal-directed actions: The role of movements and action effects. *Acta Psychologica*, 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.
- 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. (2004). Contiguity and contingency in action–effect learning. *Psychological Research*, 68, 138–154.
- 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.
- Fabbri-Destro, M., & Rizzolatti, G. (2008). Mirror neurons and mirror systems in monkeys and humans. *Physiology*, 23, 171–179.
- Falck-Ytter, T. (2008). Face inversion effects in autism: A combined looking time and pupillometric study. *Autism Research*, 1, 297–306.
- Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, 9, 878–879.
- Fawcett, C., & Gredebäck, G. (2013). Infants use social context to bind actions together into a collaborative sequence. *Developmental Science*, 16, 841–849.
- Foster, S. M., Kiskey, 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.
- Fuster, J. M. (1993). Frontal lobes. *Current Opinion in Neurobiology*, 3, 160–165.
- Fuster, J. (2002). Physiology of executive functions: The perception–action cycle. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 96–108). New York: Oxford University Press.
- Gerbing, D. W., Ahadi, S. A., & Patton, J. H. (1987). Toward a conceptualization of impulsivity: Components across the behavioral and self-report domains. *Multivariate Behavioral Research*, 22, 357–379.
- Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The one-year-olds' naive theory of rational action. *Trends in Cognitive Science*, 7(7), 287–292.
- 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, 564–574.
- 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.
- Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition*, 114, 197–206.
- 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.
- Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review*, 77, 73–99.
- 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* (Vol. 164, pp. 285–301).
- Hauf, P., & Aschersleben, G. (2008). Action–effect anticipation in infant action control. *Psychological Research*, 72, 203–210.
- 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.
- Herbert, O., & Butz, M. V. (2012). Too good to be true? Ideomotor theory from a computational perspective. *Frontiers in Psychology*, 3, 494.
- Herwig, A., & Horstmann, G. (2011). Action–effect associations revealed by eye movements. *Psychonomic Bulletin & Review*, 18, 531–537.
- 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. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, 143, 1190–1192.
- Hommel, B. (1996). The cognitive representation of action: Automatic integration of perceived action effects. *Psychological Research*, 59, 176–186.
- Hommel, B., & Elsner, B. (2009). Acquisition, representation, and control of action. In E. Morsella, J. A. Bargh, & P. M. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 371–398). New York: Oxford University Press.
- Hommel, B., Müssele, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24, 849–937.

- 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 & Review*, 19, 198–203.
- Hull, C. L. (1931). Goal attraction and directing ideas conceived as habit phenomena. *Psychological Review*, 38, 487–506.
- Hupe, J. M., Lamirel, C., & Lorenceau, J. (2009). Pupil dynamics during bistable motion perception. *Journal of Vision*, 9, 1–19.
- 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: Macmillan/Harvard University Press.
- 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.
- Kachergis, G., Wyatte, D., O'Reilly, R. C., De Kleijn, R., & Hommel, B. (2014). A continuous time neural model for sequential action. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, 369, 1655.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, New York: Prentice-Hall.
- Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, 154, 1583–1585.
- 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 & Development*, 23, 253–270.
- 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.
- 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.
- 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.
- Klapp, S. T. (1995). Motor response programming during simple and choice reaction time: The role of practice. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1015–1027.
- 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 & Development*, 29, 535–544.
- 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.
- 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.
- Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior*. New York: Wiley.
- Lewkowicz, D. J. (2004). Serial order processing in human infants and the role of multisensory redundancy. *Cognitive Processing*, 5, 113–122.
- 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., Rüniger, D., & Frensch, P. A. (2000). The role of inhibition in the regulation of sequential action. *Psychological Science*, 11, 343–347.
- 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.
- 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.
- 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.
- Marcovitch, S., & Lewkowicz, D. J. (2009). Sequence learning in infancy: The independent contributions of conditional probability and pair frequency information. *Developmental Science*, 12, 1020–1025.
- Marcus, F., Fernandes, K. J., & Johnson, S. P. (2007). Infant rule learning facilitated by speech. *Psychological Science*, 18, 387–391.
- 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.
- 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.
- McCormack, T., & Atance, C. M. (2011). Planning in young children: A review and synthesis. *Developmental Review*, 31, 1–31.
- Michel, G. F. (1998). A lateral bias in the neuropsychological functioning of human infants. *Developmental Neuropsychology*, 14, 445–469.
- Meltzoff, A. N. (2007). The “like me” framework for recognizing and becoming an intentional agent. *Acta Psychologica*, 124, 26–43.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81–87.
- Montare, A. (1977). Human temporal behavior and discrimination reversal learning. *Pavlovian Integrative Physiological and Behavioral Science*, 12, 232–246.
- Müsseler, J., & Prinz, W. (2002) (Hrsg.). *Lehrbuch Allgemeine Psychologie*. Heidelberg: Spektrum Akademischer Verlag.
- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self regulation: Advances in research* (Vol. IV, pp. 1–18). New York: Plenum.
- Olofin, 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. (2012). Action mirroring and action understanding: An ideomotor and attentional account. *Psychological Research*, 76, 760–767.
- 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).
- 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*, 7, 774–779.
- 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.
- Piaget, J. (1936, 1963). *The origins of intelligence in children*. New York: W.W. Norton & Company Inc.
- 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.
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9, 129–154.
- Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action* (pp. 167–201). Berlin: Springer.
- 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 M. Jeannerod (Ed.), *Attention and performance XIII* (pp. 321–342). Hillsdale, NJ: Erlbaum.
- Rumelhart, D. E., & Norman, D. (1982). Simulating a skilled typist: A study of skilled cognitivemotor performance. *Cognitive Science*, 6, 1–36.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70, 27–52.
- Sakai, K., Hikosaka, O., & Nakamura, K. (2004). Emergence of rhythm during motor learning. *Trends in Cognitive Sciences*, 8, 547–553.
- 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.
- 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.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools Inc.
- Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, 18, 668–671.
- 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.
- 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* (pp. 117–152). New York: Academic Press.
- Stocker, 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.
- Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- 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.
- 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., & Biro, S. (2012). Primacy of information about means selection over outcome selection in goal attribution by infants. *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. A., 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., 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.
- 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).
- Wentworth, N., Haiith, 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 cloth to retrieve a distant object. *Developmental Psychology*, 35, 651–667.
- 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. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69, 1–34.
- Woodward, A. L., & Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science*, 11, 73–76.
- Woodward, A. L., Sommerville, J. A., Gerson, S., Henderson, A. M. E., & Buresh, J. S. (2009). The emergence of intention attribution in infancy. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 51). Academic Press.
- 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.
- 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, 962–978.
- Ziessler, M., & Nattkemper, D. (2001). Learning of event sequences is based on response–effect learning: Further evidence from a serial reaction task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 595–613.