The handle http://hdl.handle.net/1887/26944 holds various files of this Leiden University dissertation

**Author:** Verschoor, Stephan Alexander  
**Title:** Learning about goals : development of action perception and action control  
**Issue Date:** 2014-06-25
Chapter 6: Where do action goals come from? Evidence for spontaneous action-effect binding in infants

Stephan Verschoor, Maaike Weidema, Szilvia Biro & Bernhard Hommel
Chapter 6

Abstract

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

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

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

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

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

The present study sought for more direct evidence for the acquisition of bidirectional action-effect in infancy. Although we hypothesize that the ability to learn action-effect associations is present from early infancy, action production and action control are extremely limited in very young babies, thus rendering behavioural measurements difficult. Since investigating bidirectional action-effect acquisition requires participants to carry out some sort of motor actions, a possible lack of evidence in very young infants could simply be due to limitations in their motor capabilities rather than to the inability to associate and rely on action-effect relations. Adult and children studies on learning action-effect associations typically used manual actions such as pressing a button. Good control over these types of manual actions develops in the second half of the first year (Belsky & Most, 1980). In addition, previous research showed that the earliest age at which infants' own actions are influenced by observed effects of others' actions is around 9 months (Hauf & Aschersleben, 2008). At this same age infants begin to distinguish means from ends in their own behaviour (Piaget, 1936; Goubet, Rochat, Maire-Leblond, & Poss, 2006; Willatts, 1999). This ability is often thought to be a prerequisite for goal-directed action control, as it enables infants to specify and represent goals before performing the corresponding actions.
(Hauf, 2007). Therefore, we expected that at 9 months infants will likely show evidence for bidirectional action-effect acquisition via their own manual exploratory behaviour. Furthermore, Elsner and Aschersleben's study (2003) suggests that there are developmental changes in the effect that observed action-effect relations have on infants' own behavior. They found that with a higher task demand the behavior of 12 but not 9 months old infants was influenced by an observed action-effect. In addition, a further step was made at around 15-18 months when infants' behavior started to be affected by whether the effect they could produce matched with the observed effect. Testing 9, 12 and 18 months olds' ability to acquire bidirectional action-effects by their own exploratory behavior can help clarify whether the nature of these changes lies in the development of specific imitative skills for transferring others action into one's own action or in the accuracy of encoding specific action-effect relations due to, for example, increasing working memory capacity or brain maturation (Diamond, 2006).

Our task was modelled after the experimental setup used in the free-choice experiments of Elsner and Hommel (2001). Elsner and Hommel had adults carry out self-chosen left and right key presses in response to a visual trigger stimulus. Each key press produced a particular sound (e.g., left key → low tone, right key → high tone), and even though these sounds were irrelevant for the task, it was assumed that participants would acquire bidirectional associations between key presses and tone representations. After this acquisition phase, participants were again freely choosing left and right key presses, but in this test phase the visual trigger was replaced by an auditory trigger stimulus: high and low tones (identical to the previous action effects) presented in random sequence. As predicted, people were quicker and more likely to choose the action that previously had produced the currently presented trigger tone (e.g., they were quicker and more likely to press the left key when hearing the low than the high tone), suggesting that key presses and tones were indeed associated in a bidirectional fashion.
When adopting this paradigm for our infant subjects, we realized after piloting that versions of the original binary-choice task were too demanding. We therefore simplified the task in such a way that only one (very large) touch-sensitive key was presented to the infants, the two action alternatives being the touching or not touching of the key. Moreover, given Dutzi and Hommel’s (2009) demonstration of action-effect learning after just one presentation, we also greatly reduced the number of acquisition- and test trials, so to minimize demands on infants’ limited attention. In the acquisition phase, infants were presented with two multimodal (and thus very salient and attention-grabbing) events, one being self-produced by touching the key (a true action effect) and another that was presented while infants were prevented from touching the key, so that the event was not self-produced. In the test phase, we presented the two multimodal events in different blocks and expected the previously self-produced event stimulating quicker and more spontaneous key touches than the action-independent event.

**Materials and methods**

**Participants**

Three groups of infants were tested: 22 9-month-olds (mean: 9.09 months, SD= .18, 7 female), 21 12-month-olds (mean: 12.10 months, SD= .26, 10 female), and 22 18-month-olds (mean: 18.02 months, SD= .32, 8 female). Four additional 9-month-olds, three 12-month-olds, and two 18-month-olds were excluded due to fussiness. One 12-month-old was excluded due to experimental error. The participants were recruited through advertisements in local papers, daycares, maternity wards, and general practitioners. Infants were randomly and equally distributed across conditions, and they received small gifts as reward. An informed consent was obtained from all caretakers.
Chapter 6

Apparatus

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

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

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

---

12 We decided not to balance action-independent and self-produced effects across participants after a pilot experiment had revealed no preference for one or the other effect. The pilot investigated 15 participants, five from each of the three age groups. Infants were shown the button but prevented from touching; and it was checked whether the infant was actually looking at the button. We then presented the two bimodal events in random order, and the infant had the opportunity to touch the button for 30 seconds. A short break, during which infants were prevented from reacting again, divided the two presentations. The two effect stimuli did not yield any reliable difference in the latencies (F(1,12)=.627,
After the acquisition phase there was a 30-second break during which caretaker and infant turned away from the monitor while engaging in entertaining the infant. The test phase consisted of three trials, in which the previously self-produced event, the previously experienced action-independent event, or no event was presented. Each of these trials lasted 30 seconds, during which the infant could freely touch the key. In between the test trials there was a 10-sec break during which the caretaker prevented the infant from responding by holding his/her hands. The self-produced and action-independent events were presented first and second, or vice versa, with the order being balanced across participants, while the baseline trial was always administered last—so to minimize possible forgetting and extinction of action-effect associations. Methods were approved by the ethical committee of the Leiden University, Institute for Psychological Research.

**Data acquisition**

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

\[ p=.444 \]

or the number of responses \( F(1,13)=.398, p=.389 \), nor was there any hint to interactions with age or order of presentation.
frequency was calculated for each infant based on his/her total number of responses in all three test trials, thus correcting for motivational differences (see Table 1).

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

| Test-Trial Type | Self Produced | Action Independent | No Event |
|----------------|---------------|--------------------|---------|
| Age groups     | #  | %       | #  | %       | #  | %       |
| 9-Month olds   | 5.00| 38.95  | 4.91| 41.23  | 2.50| 19.73  |
| 12-Month olds  | 3.67| 36.38  | 4.33| 41.60  | 2.24| 22.02  |
| 18-Month olds  | 4.86| 49.53  | 4.36| 26.67  | 2.41| 23.75  |
Chapter 6

Results

Acquisition phase

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

Test phase

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

A between-subjects ANCOVA was carried out on the latencies of first responses in the test trials, with trial type (previously self-produced vs. action-independent stimulus event) as within-subjects and age group (9, 12, 18 months) as between-subjects factor (see Figure 2). In this analysis we considered the number of undetected responses during acquisition as a covariate, so to control for extinction, and the total number of responses during all three test trials, so to control for the overall level of activity. We found a significant main effect of trial type, $F(1,51)=4.67$, $p = .035$, $\eta^2_p = .564$: Infants responded significantly faster to self-produced events (5.57 s) than to action-independent events (6.64 s). No effects of age group or interactions with the covariates were found.
Figure 2. Mean Latencies (+ SE) for 9-month-olds (n = 22), 12-month-olds (n = 21) and 18-month-olds (n = 22) in associated and non-associated test-trials.
We ran another mixed-factor ANCOVA on the response proportions in the test trials, normalized through an Arcsin transformation. Trial type and age group served as factors (see Figure 3) and the number of undetected responses during the acquisition as covariate. There was no main effect of trial type ($F(1,61)=0.70$, $p = .41$, $\eta^2 p = .011$) but a close-to-significant interaction of trial type and age group, $F(2,61) = 2.89$, $p = .063$, $\eta^2 p = .087$. Separate exploratory between-subjects ANCOVAs revealed that there was no trial-type effect in the 9- or the 12-month-olds, $F$s<1, but responses were more frequent to self-produced than action-independent events in the 18-month-olds, $F(1,20) = 5.31$, $p = .032$, $\eta^2 p = .210$.

![Figure 3. Mean Percentage of responses (+ SE) for 9-month-olds (n = 22), 12-month-olds (n = 21) and 18-month-olds (n = 22) in associated and non-associated test-trials.](image-url)
Chapter 6

Discussion

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

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

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

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