Developmental Science 10:3 (2007), pp 379–398 DOI: 10.1111/j.1467-7687.2006.00544.x

PAPER

# Infants' perception of goal-directed actions: development through cue-based bootstrapping

**Szilvia Biro[1] and Alan M. Leslie[2]**

1. *Department of Psychology, Leiden University, The Netherlands*
2. *Department of Psychology and Center for Cognitive Science, Rutgers University, USA*

## Abstract

*It is now widely accepted that sensitivity to goal-directed actions emerges during the first year of life. However, controversy still surrounds the question of how this sensitivity emerges and develops. One set of views emphasizes the role of observing behavioral cues, while another emphasizes the role of experience with producing own action. In a series of four experiments we contrast these two views. In Experiment 1, it was shown that infants as young as 6 months old can interpret an unfamiliar human action as goal-directed when the action involves equifinal variations. Experiments 2 and 3 demonstrated that 12- and 9-month-olds are also able to attribute goals to an inanimate action if it displays behavioral cues such as self-propelledness and an action-effect. In Experiment 4, we found that even 6-months-olds can encode the goal object of an inanimate action if all three cues, equifinality, self-propelledness and an action-effect, were present. These findings suggest that the ability to ascribe goal-directedness does not necessarily emerge from hands-on experience with particular actions and that it is independent from the specific appearance of the actor as long as sufficient behavioral cues are available. We propose a cue-based bootstrapping model in which an initial sensitivity to behavioral cues leads to learning about further cues. The further cues in turn inform about different kinds of goal-directed agents and about different types of actions. By uniting an innate base with a learning process, cue-based bootstrapping can help reconcile divergent views on the emergence of infants' ability to understand actions as goal-directed.*

## Introduction

The nature and the emergence of infants' ability to understand actions as goal-directed has become a central issue in cognitive infancy research in the last decade. When, how and under what conditions are infants willing to go beyond the spatio-temporal pattern of observed actions and interpret them as performed in order to achieve an end-state? Several theorists assumed (e.g. Dennett, 1987; Leslie, 1994; Tomasello, 1999; Csibra & Gergely, 1998) that giving meaning to an action by attaching a goal to it is one of the important prerequisites for 'theory of mind' (interpreting and predicting other people's actions in terms of intentional mental states). This paper describes and contrasts two theoretical approaches to the origins of goal-directed action interpretation (the cue-based and the experience-based accounts) and provides empirical evidence in favor of the cue-based account.

## Cue-based theories

One group of theories credits the infant with innately based, domain-specific and/or modular systems that are sensitive to certain behavioral cues (such as self-propelledness, direction of movement, equifinal variation of the action, contingent reactivity, and efficiency of the action toward the goal) in identifying agents and goal-directedness (Premack, 1990; Leslie, 1994, 1995; Baron-Cohen, 1994; Jonhson, Slaughter & Carey, 1998; Csibra & Gergely, 1998; Gergely & Csibra, 2003). While these theories differ from each other in many respects, they all suggest that ascribing goal-directedness to an action is initially independent from the perceptual appearance of the actor. Thus, infants are assumed to be able to apply goal-directed action interpretation to a wide range of entities including unfamiliar inanimate agents without any human features and this ability does not require extensive experience with human actions.

Address for correspondence: Alan M. Leslie, Department of Psychology and Center for Cognitive Science, Rutgers University, 152 Frelinghuysen Road, Piscataway, NJ 08854, USA; e-mail: aleslie@ruccs.rutgers.edu or Szilvia Biro, Department of Psychology, Cognitive Psychology Unit, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands; e-mail: sbiro@fsw.leidenuniv.nl

One set of views about the role of abstract behavioral cues stresses the efficiency of *single* cues, for example, Premack (1990) and Baron-Cohen (1994) who claimed that infants automatically interpret the action of an agent as intentional and goal-directed as long as it exhibits self-propelled movement. Other views stress the *multiplicity* of cues that can contribute probabilistically to agency judgments. For example, Leslie (1994, 1995) proposed a tripartite hierarchical modular mechanism, which on its second level (theory of mind mechanism, system 1) deals with the actional properties of agents such as attaining goals and reacting to the environment. This second subsystem receives its input from the first subsystem (the 'theory of body mechanism'), which can categorize objects as 'physical or mechanical agents' with an internal and renewable source of 'energy' on the basis of such behavioral cues as self-propulsion. The second level can consider circumstances that are distant in time and space from the agent and represent the agent's behavior in relation to spatially distant features and future/possible states of affairs. Thus, the outcome of the agent's action is a useful cue at this level as it can be teleologically interpreted as the goal state of affairs.

In a similar vein Csibra and Gergely (1998) proposed that infants are equipped with an inferential system, which they call the 'teleological stance', for generating non-mentalistic goal-directed action representations. This system establishes a specific relation among three representational elements: the observed behavior (the action), a possible future state (the goal), and the relevant aspects of the reality (constraints of the action). This representational structure forms a teleological representation only if it satisfies the 'principle of rational action' that states that an action can be explained by a goal-state if it appears as the most efficient action towards the goal-state that is available within the constraints of reality. Csibra and his colleagues (Csibra, Biro, Koós & Gergely, 2003) argue that the abstract behavioral cues that are critical in mediating the attribution of goal-directedness can be derived from the 'principle of rational action' itself: the agent's ability of equifinal variations of the action, and that the outcome of the action involves a change in the environment.

What kind of experimental evidence is available to support these theoretical proposals concerning the sensitivity to abstract behavioral cues in generating goal representation irrespective of the appearance of the actor? Heider and Simmel's classic work (1944) showed that adults spontaneously provide rich intentional and goal-directed action interpretations when all they see are animated interactions of triangles and circles. Heider and Simmel's study was replicated and their findings found to be very robust and consistent across a wide range of cultures (see Scholl & Tremoulet, 2000). Castelli and her colleagues (Castelli, Frith, Happe & Frith, 2002) found that even high-functioning adults with autism can give a goal-directed description to a moving triangle on a screen. Other contemporary studies (see Tremoulet & Feldman, 2000; Scholl & Tremoulet, 2000, for a review) have attempted to pinpoint the specific motion cues responsible for the perception of animacy or intentionality (including goal-directedness) in adults in simple animations with varying success, on the assumption that it involves a fast, automatic, irresistible process similar to the perception of physical causality (Michotte, 1946). Berry and Springer (1993) found that 3- and 4-year-old children give similar intentional descriptions to Heider and Simmel's type of events of two-dimensional geometric figures as do adults. Furthermore, Montgomery and Montgomery (1999) showed that 3-year-old children can judge if the action of a ball was accidental or intentional purely on the basis of motion cues, namely whether the behavior of the ball appeared to be persistent and was adjusted toward its goal.

Recent research with infants also demonstrates sensitivity to behavioral cues exhibited by geometric figures. Rochat and his colleagues (Rochat, Morgan & Carpenter, 1997) found that 3-month-olds looked differently at two discs, depending on whether they were moving independently or in a systematic interaction ('chased' each other) while the relative spatial-temporal and other dynamic features were equal in both cases. Schlottman and Surian (1999) showed that 9-month-old infants understand that an agent can react to another without contact: in their visual habituation study infants showed sensitivity to the temporal proximity of the movement of two squares exhibiting self-propulsion and non-rigid transformation. Johnson, Slaughter and Carey (1998) showed that 12-month-old infants follow the spatial orientation of a faceless, inanimate and novel agent that exhibits contingent reactivity to the infant's behavior. All of these findings suggest that the specific appearance of the actor is not critical for infants' judgments of agency.

More direct empirical evidence for relying on abstract behavioral cues in interpreting actions as goal-directed comes from a series of visual habituation studies by Gergely and his colleagues (Gergely, Nádasdy, Csibra & Biro, 1995) and Csibra and his colleagues (Csibra, Gergely, Biro, Koós & Brockbank, 1999; Csibra *et al.*, 2003) who demonstrated the presence of teleological action representation in 9- and 12-month-old infants. In one of the experiments using computer animation of 2D geometric figures, infants were habituated to a visual event in which they observed a small circle repeatedly approach and make contact with a large circle by 'jumping over' a rectangular figure separating them. Adults typically interpret this behavior as a goal-directed action. In the test trials the rectangular figure (the 'obstacle') was removed. Infants saw either a novel action (the small circle approached the large circle in a straight line) that was the most efficient action toward the goal in the changed circumstances, or the already familiar jumping action which, however, was no longer the most efficient action to achieve the same goal-state. Infants showed less recovery of attention to the novel straight-line action, which indicates that they interpreted the action in the habituation events as goal-directed and predicted the most efficient action to achieve the inferred goal in the changed situation. Similar experiments by Csibra and his colleagues (2003) demonstrated that by 12 months infants could also use the 'principle of efficient (rational) action' for making productive inferences about unseen aspects of goal-directed actions (e.g. unseen goals and occluded obstacles). Furthermore, Shimizu and Johnson (2004) found that 12-month-old infants are able to attribute goals to an unfamiliar non-human 3D object that responds contingently and can vary its action. Finally, the study of Kuhlmeier, Wynn and Bloom (2003) indicates that 12-month-olds may be able to attribute dispositional states to geometric figures and use these to predict their future actions.

## Experience-based theories

Now we turn to another class of theories on the emergence of infants' goal-directed action interpretation. This approach proposes that infants' understanding of actions as goal-directed is, from the start, restricted only to human actions and therefore tied to specific appearances. Furthermore, it is assumed that this understanding is acquired gradually through experiences with particular actions.

Meltzoff argues (e.g. Meltzoff & Moore, 1994) on the basis of his extensive imitation research that very young infants, even newborns, have the ability to identify human actions through biomechanical bodily movements (such as facial expressions). He claims that infants only consider actions that they categorized as human as pursuing goals. Meltzoff (1995) supports his claim by demonstrating that 18-month-old infants, after having observed an adult perform three failed attempts to achieve a goal, re-enact the intended act (that they have never seen realized) and not the failed attempts. On the other hand, when the same (unsuccessful) actions were modeled not by a person but by a mechanical device, the infants failed to produce the target act. This finding, however, was challenged by the study of Johnson, Booth and O'Hearn (2001) in which Meltzoff's experiment was replicated with 15-month-old infants using a non-human agent with a face that contingently interacted with the experimenter and the infant. The authors found that infants were able to infer and re-enact the goal of the failed action and suggested that goal re-enactment is 'mediated by a concept that is not isomorphic with the concept person' (p. 637).

The strongest claim for the experience-based account of goal-directed action interpretations comes from Amanda Woodward and her colleagues' recent work with young infants. Using a visual habituation method, Woodward (1998) presented 6- and 9-month-old infants with an action in which a hand repeatedly reached towards and grasped one of two toys sitting on a stage. After the infants were habituated to this event, the experimenter swapped the two toys behind a screen. In the test phase the hand either grasped the same toy as before, which, however, was at a new location, or the other toy at the same location. Looking times for these two events were markedly different: both 6- and 9-month-old infants looked longer if the hand grasped the new toy at the old location than if it grasped the old toy at the new location. This result indicates that infants associated the grasping hand with the grasped object rather than with its location, that is, they expected the hand to reach toward the same toy. In other versions of the same experiment, instead of a human hand, inanimate novel objects (such as a rod, a mechanical claw or a flat occluder) were used to touch or grasp the target toy. Woodward found that in these conditions 6- and 9-month-old infants looked equally long at the two test events (or longer in old toy/new location event), suggesting that the infants did not selectively encode the goal object in the case of inanimate actors. Woodward argued that goal-directed interpretations are specific to human actors, and that 'by 6 months infants already have begun to draw a line between animate and inanimate entities, interpreting only motions of the former, but not the latter in terms of the relation between agent and object' (Woodward *et al.*, 2001, p. 155).

To further specify the range of actions infants would consider in terms of goals, Woodward (1999) applied the same set-up using an unfamiliar human action: the back of the experimenter's hand dropped on the target toy. Six- and 9-month-old infants showed no preference toward any of the test events. In another study (Guajardo & Woodward, 2004), 7- and 12-month-old infants were shown a grasping hand whose surface properties were, however, obscured by a glove. Neither of the two age groups was able to interpret the grasping action of the unfamiliar gloved hand in terms of goals unless they had seen that the gloved hand belonged to a person. Other studies have looked into the connection between the production and understanding of goal-directed actions. For example, a correlation was found between the ability to identify the goal of an observed means–end sequence and executing a similar means–end task in 12-month-olds (Sommerwille & Woodward, 2005) and between infants' own pointing and their point comprehension at 9.5 months (Woodward & Guajardo, 2002). Finally, Sommerville and her colleagues (Sommerville, Woodward & Needham, 2005) showed that if infants as young as 3 months old are allowed to practice making contact and picking up toys by using a mitten that is covered with Velcro fabric on the palm, they can attribute a goal when they observe a grasping action of a hand wearing a similar mitten. On the basis of these results Woodward and her colleagues argued that initially infants do not interpret an overly broad set of behaviors as goal-directed; instead, they gradually construct such interpretations with specific actions. Understanding goal-directedness derives from experience with particular actions (from their own actions and from interactions with social partners), rather than being a product of a general mechanism that is sensitive to the motion properties of any actions.

## Is there a real conflict?

At first glance it appears that Woodward and her colleagues' results seriously challenge the cue-based account in terms of the time of onset of goal-directed interpretations, the range of entities to which infants are willing to attribute goals, and the role that abstract behavioral cues play in generating such interpretations. However, along with others (e.g. Kiraly, Jovanovic, Prinz, Aschersleben & Gergely, 2003) we believe that this challenge can be deflected. One reason infants might not selectively associate the action of the inanimate objects (e.g. a mechanical claw) and the unfamiliar human actions (e.g. 'dropping the back of the hand' or the 'grasping gloved hand') with the target object in Woodward's experiment, is that the actions were impoverished in the behavioral cues infants are most sensitive to in evaluating the goal-directedness of an action. For example, in the beginning of each habituation trial the actor entered the stage from behind a curtain and, after touching or grasping the target toy, stayed motionless for the rest of the trial exhibiting no other behavioral cues. Leslie (1994, 1995) suggested that judgments of mechanical agency tacitly reflect the likelihood of the agent moving through an internal energy source. Furthermore, these judgments take conditional probabilities into account, so that periods of non-motion are considered too as negative evidence. Thus, the mere repetition of the same action together with a stationary period across trials will not provide a sufficient cue for agency/goal-directedness. Therefore, it is possible that in a modified version of Woodward's set-up in which unambiguous behavioral cues are present, infants would be able to ascribe goal-directedness to inanimate or unfamiliar human actions.

One could argue that Woodward's grasping hand version also lacked such cues, yet 6-month-old infants were able to generate a goal-directed action interpretation. As Woodward herself acknowledges, however, grasping is an action with which infants have ample experience by 6 months of age. Kiraly and her colleagues (2003) point out that in natural circumstances the experience that infants have with grasping actions is in fact very rich in behavioral cues. Infants can adjust their grasping action according to the differences in weight, shape or texture of the target objects. Von Hofsten (1983, 1991; Von Hofsten & Fazel-Zandy, 1984) showed that very young infants are able to co-ordinate their grasping action with the future position of a moving target. Furthermore, a grasping action is typically followed by a change in the state of the grasped object. The object is moved, picked up or thrown away by the infants and similar outcomes occur when infants observe the effect of others' grasping actions. Indeed, Leslie (1982, 1984) showed that 4- and 6-month-olds perceive a causal structure when a hand grasps and picks up an object. Thus, grasping is strongly associated with behavioral cues such as self-propelledness, equifinal variations, and mechanical effects on the environment. Watching the hand grasping the toy in Woodward's experiment may have triggered the memory of these associated properties, which could then have enabled infants to interpret the action as goal-directed despite the absence of direct perceptual evidence.

Kiraly and her colleagues (2003) and Jovanovic and her colleagues (Jovanovic, Király, Elsner, Gergely, Prinz & Aschersleben, under review) recently tested the hypothesis that 'adding' a behavioural cue would allow young infants to interpret an unfamiliar human action as goal-directed. They designed a modified version of Woodward's 'dropping the back of the hand' action by introducing a salient action-effect: after the hand contacted the toy it pushed the toy away toward the back of the stage. Both studies found that with the help of this cue 6-, 8- and 10-month-olds were able to attribute a goal to the otherwise unfamiliar action. These results suggest that, as the cue-based approach claimed, infants do not necessarily need to have experience with a particular action to be able to interpret it as goal-directed as long as the action exhibits the necessary behavioral cues.

In order to assess the effectiveness of multiple behavioral cues, it needs to be shown that other cues can also trigger goal-directed action interpretations of unfamiliar human actions. Furthermore, in order to test the other major assumption of the cue-based account, namely, that ascribing goal-directedness is not restricted to human action, it must be demonstrated that the availability of similar behavioral cues will allow young infants to consider not only animate actions as goal-directed, but inanimate actions as well. In the first experiment we tested the hypothesis that young infants can apply goal-directed action interpretation to an inanimate action as well as to an unfamiliar human action when these actions exhibit equifinal variation of behavior. Heider (1958) argued that the equifinal structure of an action helps identify the actor's goal. That is, as environmental conditions vary, differing motion patterns used to bring about one and the same consequence reveal the actor's goal. Recent studies suggest that this cue may not only play a critical part in goal attribution in adults, but also in infants (e.g. Gergely *et al.*, 1995). However, its role has not yet been tested directly.

## Experiment 1

We applied Woodward's (1998) set-up with an animate (hand) and an inanimate actor (a paper tube) performing the same action: rhythmic poking of the target object. The poking action was carried out in a manner that indicated that the actor is capable of equifinal variations of behavior (the actor poked from different angles and the target toy was touched at different spots). Infants typically do not poke (that is, explore toys with extended index finger) until they are about 9 to 12 months old (Blake, O'Rourke & Borzellino, 1994; Bates, Benigni, Bretherton, Camaioni & Volterra, 1979). They might have seen adults performing somewhat similar poking actions, but not in the particular way that was shown to them in the experiment – repetitively from different angles. Therefore, the action used in the experiment is at least relatively unfamiliar at 12 months and for infants younger than 9 months highly unfamiliar.

Could infants mistake poking for pointing? We believe this is unlikely. Infants start pointing between 9 and 12 months according to Bates *et al.* (1979) but only around 12 months according to extensive data from Butterworth *et al.* (Butterwoth, Franco, McKenzie, Graupner & Todd, 2002). Index finger extensions occur in very young infants (e.g. Trevarthen, 1977; Hannan, 1987) but not in relation to an object. Thus, 6-month-olds will have no experience of their own pointing. Infants of this age will have experience of observing pointing in adults but again never in the manner displayed in this experiment – repetitively, from different angles, and, most importantly in this context, making contact with the object. Whether a poking action can be considered as an action that is part of pointing is less clear. It is assumed by some (e.g. Bates *et al.*, 1979) that poking might be a precursor for pointing (as pre-pointing), but others (e.g. Blake *et al.*, 1994) found that they emerge simultaneously. Although the two actions are somewhat similar superficially, the major difference is the lack of contact with the object in pointing. Furthermore, from the outset pointing is a referential action (Butterworth *et al.*, 2002), whereas poking is a goal-directed action. Some have argued that understanding these two types of action develops from distinct mechanisms (Csibra, 2003).

Infants commonly observe others reaching and touching objects. Although there are again superficial similarities between these actions and poking, our question is whether infants interpret a given action as goal-directed. Indeed, it can be argued that goal-directedness is what turns motion into action (e.g. Dennett, 1987; Leslie, 1994; Tomasello, 1999; Csibra & Gergely, 1998). If experience with touching objects or reaching for an object were sufficient, then there would be evidence for goal-directed action interpretation for any observed hand action that involves some form of touching or reaching for an object. However, this is not the case. Woodward has found that actions other than the grasping action, for example, touching the object with the back of the palm or touching the object with the index finger once, did not elicit goal-attribution in 6- or 9-month-old infants (Woodward, 1999; Woodward & Guajardo, 2002).

Repetitive poking from different angles will therefore be an unfamiliar action for our youngest babies and for our oldest babies relatively unfamiliar (compared to grasping, for example). We hypothesized that nevertheless infants will encode the goal of both the animate and inanimate actions and that looking times across these conditions would not differ. Three age groups were tested: 12-, 9-, and 6-month-olds.

### *Method*

#### Participants

Forty-eight 12-month-old (24 males and 24 females, mean age = 12 months 1 day, SD = 15.9 days, range 11 months to 12 months and 29 days), 48 9-month-old (28 males and 20 females, mean age = 8 months 28 days, SD = 9.4 days, range from 8 months and 14 days to 9 months and 15 days), and 48 6-month-old (28 males and 20 females, mean age = 6 months 8 days, SD = 14.2 days, range from 5 months and 2 days to 6 months 29 days) infants participated in the study. An additional 13 12-month-old, seven 9-month-old, and 13 6-month-old infants were also tested but were excluded from the data analysis because of fussiness (five, four and six, respectively), experimental error (six, two and three), inattentiveness during familiarization (two, one and three), or looking less than 2 seconds on each of the test trials (zero, zero and one) (see Procedure below). All of the subjects in this and the subsequent experiments were healthy, full-term infants living in the New Brunswick, NJ area and were recruited through mailings.

### Material/apparatus

Infants sat in their parent's lap in a darkened experimental room (a curtained booth) facing a stage from a distance of approx. 1 meter. The stage was 70 cm high × 103 cm wide × 48 cm deep and was made from white cardboard. A black curtain could be raised to block the stage from view between trials. An opening on the right side of the stage, which was covered by a white curtain, allowed the experimenter to enter her arm or an object into the stage area. Two concealed lamps illuminated the stage. A computer signal turned the lights on to start the first trial. A video camera focusing on the baby's face was mounted above the stage peeping through the opening of a dark blue curtain. All infants were videotaped and a hidden observer, blind to the experimental condition, timed the infants' looks in the test trials at the displays from a black and white head-and-shoulders en face video image. The observer could neither see the stage area nor the experimenter's arm. Another camera mounted above the head of the infant recorded the stage area. A split screen recording was made, with the stage image on the top and the baby's face on the bottom part of the display. The timing of the trials and the movements of the experimenter were regulated by a metronome beating softly every second.

### Stimuli and procedure

*Familiarization trials.* Half of the infants in each age group were assigned to the Poking Hand condition, while the other half participated in the Poking Tube condition. In both conditions infants saw four familiarization trials. At the start of each trial the curtain was lowered to reveal the stage. On the stage floor there were two toys: a yellow plastic bear and a green ball with white stripes (both were about 12 cm tall). Each toy was sitting on pedestals made from white cardboard (12 × 12 × 12 cm) which were 20 cm apart. In the Poking Hand condition, after a computer signal turned on the lights, the experimenter reached in through the opening on the right side of the stage and started to 'poke' one of the toys. In each trial she poked the toys 10 times with her extended index finger (each poke took about 2 sec including the approaching, touching and retrieving of the hand). Both the angle of the poking and the spot on the toy that was touched were varied (see Figure 1). The experimenter's arm was bare and she wore no rings or other jewelry. The arm was the only part of her that was visible. Each familiarization trial lasted about 20 seconds. Between the trials the curtain was raised and there was an approximately 3-sec pause.

The familiarization procedure in the Poking Tube condition was identical to that in Poking Hand condition with the exception that the poking action was carried out by a paper tube. The paper tube was made of two pieces of rolled-up brown cardboard paper. The main piece was 56 cm long and 4–5 cm in diameter. The smaller 'finger' piece, which was attached to the end of the main part, was 6 cm long and 2 cm in diameter (see Figure 1). Note that the paper tube did not resemble a hand; it only had similar dimensions and color. The paper tube poked the target toys in the same manner and rhythm as the hand did. The experimenter who operated the paper tube was neither visible nor audible.

In both conditions the position of the toys (left/right), the location of the poking action (near side/far side) and the target object (bear or ball) were varied and counterbalanced during the familiarization trials.

*Test trials.* After the fourth familiarization trial the curtain was raised and the positions of the two toys were switched. Then the curtain was lowered and the infants were shown the toys in their changed positions without the presence of the arm or the tube for 5 seconds. Following this short trial, one of two types of test events was given to the infants. Half of the infants saw two identical tests trials in which the arm (in the Poking Hand condition) or the paper tube (in the Poking Tube condition) was poking the same toy as had been poked during the familiarization. However, since the toys had been swapped, the poking action took place at a different location (old goal/new location test event; see Figure 1). The other half of the infants saw two tests trials in which the experimenter's arm/the paper tube poked a new toy while the action took place at the old location (new toy/old location test event; see Figure 1). Thus, there were 12–12 infants for each test trial type (old goal/new location and new goal/old location) in both conditions (Poking Hand/Poking Tube) with girls and boys distributed approximately evenly across the test events.

The online observer started to measure the looking times when the poking action started and the baby was looking at the stage. The experimenter or the paper tube continued poking in the same rhythm and manner as in the familiarization trials for the duration of the test trial. The test trial ended when infants had looked continuously

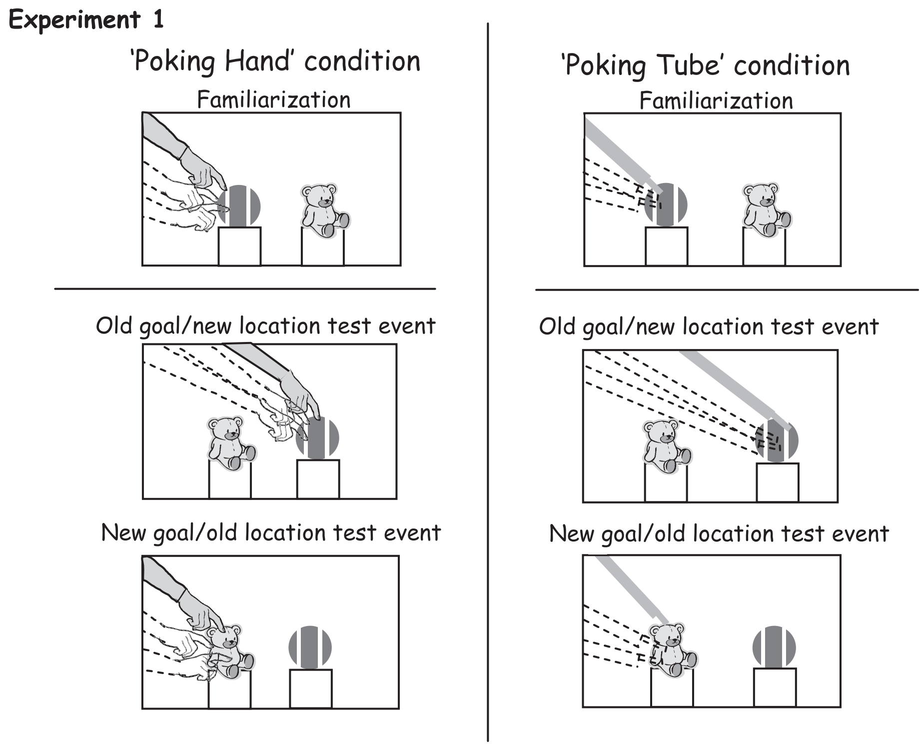


**Figure 1** *Stimulus events of the Poking Hand and the Poking Tube conditions in Experiment 1.*

for a minimum of 2 sec and then looked away continuously for 2 sec. If infants looked away before 2 sec elapsed the trial was started again. There was a pause of about 3 sec between the two test trials. Parents were asked to close their eyes during the test trials and instructed to refrain from talking to the infant during the whole session except for giving comfort when necessary.

*Reliability coding.* To assess reliability, offline observers measured each infant's looking times from the videotaped record. The observers were undergraduate students, trained in coding looking times, who were unaware of the conditions of the trials and the experimental design. Inter-observer agreement was computed for each infant's looking times in the two test trials and the agreement was accepted if it was above or equal to 95%. If the agreement was below 95%, a second offline observer was asked to measure the looking times again independently. If this did not lead to agreement with one of the other observers, then the looking times of that infant were excluded from the analyses (this happened for one participant).

### *Results*

The looking times of the two test trials were analyzed. Preliminary analysis did not reveal any effect of sex, location of the poking action (near side vs. far side) and target object (bear vs. ball). These factors therefore were omitted from further analyses. Using Condition (2) [Poking Hand vs. Poking Tube], Trial Type (2) [new goal/old location vs. old goal/new location], Position (2) [left vs. right], Age group (3) [6-, 9-, 12-month-olds] as between-subject factors and Test Trial (2) as within-subject factor an ANOVA was carried out that showed the following effects. A main effect of Trial Type was found [$F(1, 120) = 9.70$, $p < .002$]. Effect size was calculated using $\eta^2 \times 100$ which showed that Trial Type factor accounted for 7.5% of the variance. This main effect indicates that overall infants across all age groups and conditions looked longer in the new goal/old location test event than in the old goal/new location test event. The ANOVA also revealed a Position main effect [$F(1, 120) = 9.83$, $p < .002$, effect size: 7.6%] indicating that overall infants looked longer when the bear was on the right side of the stage in the test trials. Furthermore, the ANOVA showed a Test Trial [$F(1, 120) = 29.90$, $p < .0001$, effect size: 20%] main effect indicating that looking times significantly declined across the two test trials. Finally, an interaction between Test Trial and Condition was also found [$F(1, 120) = 7.18$, $p < .008$, effect size: 5.6%] which revealed that infants looked significantly longer in the new goal/old location test event than in the old goal/new location

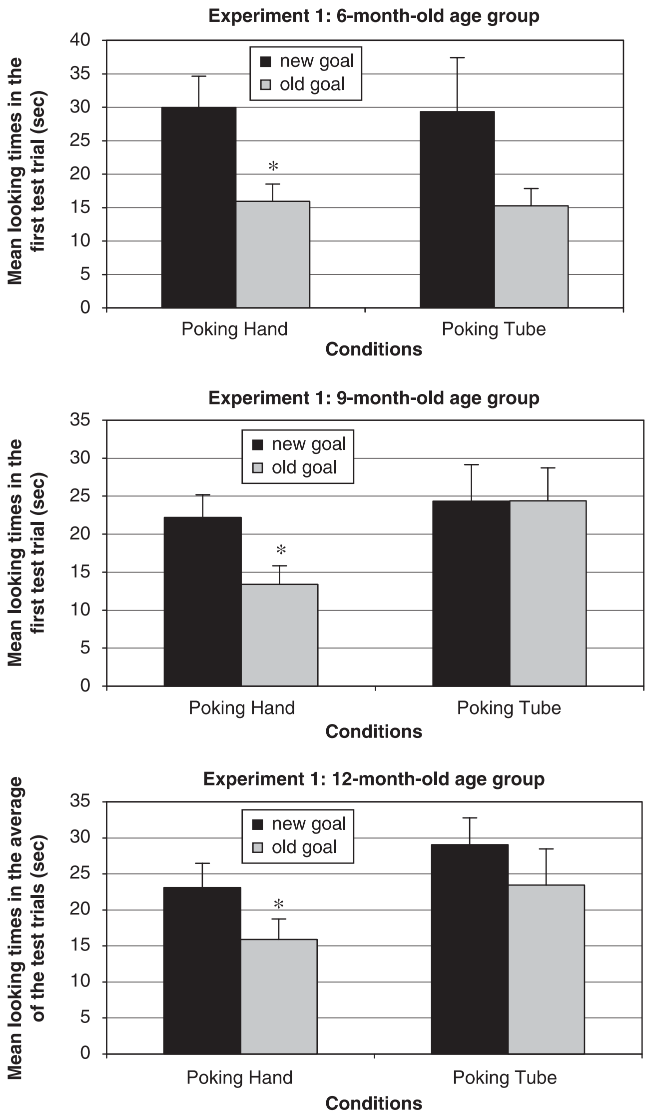


**Figure 2** *Mean looking times (bars show SEM) in the first test trial in Experiment 1 as a function of Trial Type and Age group (* =* $p < .05$*).*

test event only in the first trial [$t(142) = 3.376$, $p < .001$], but not in the second trial.

Although the ANOVA did not reveal a main effect of Age group or an interaction effect between the Age group and Condition or Trial type factors, separate $t$-tests were carried out for both conditions in each age group. Mean looking times with standard errors by age groups and conditions are shown in Figure 2. $T$-tests showed that, in the Poking Hand condition, the 6-months-old [$t(22) = 2.61$, $p < .02$, two-tailed] and the 9-months-old age groups [$t(22) = 2.27$, $p < .04$, two-tailed] in the first test trial, and the 12-months-old age group in the average looking time of the two test trials [$t(22) = 2.13$, $p < .05$, two-tailed] looked significantly longer in the new goal/old location test event than in the old goal/new location test event. In the Poking Tube condition, the 6- and 12-months-old age groups looked considerably longer in the new goal/old location than in the old goal/new location test event; however, the differences were not significant.

A follow-up analysis was carried out to test an alternative explanation for the looking pattern found for the two test events. One could argue that infants looked longer in the 'new goal/old location' test event simply because their attention was drawn (by the poking hand or tube) to a 'novel object'. This explanation was termed by Woodward (1998; see also Heineman-Pieper & Woodward, 2003) the 'spotlight effect'. Note, however, that although the spotlight effect could possibly have been present in both conditions, it was only in the Poking Hand condition that infants looked significantly longer in the new goal/old location test event than in the old goal/new location test event. One could argue though, that the poking hand – being inherently more interesting – might have been a better attention getter than the poking tube. This possibility was tested by comparing infants' attention allocated to the two objects in the two conditions. If infants in the Poking Hand condition spent more time looking at the 'novel object' than in the Tube condition, then the 'spotlight effect' could be responsible for the findings. Eighteen infants in each condition (six infants in each age group) were randomly selected and their videotapes were re-coded. For each test trial the proportion of the total trial time spent looking at the object that was poked and at the other object was calculated. Mean proportions for both conditions are depicted on Figure 3. To assess the effectiveness of the hand and the tube as spotlights, for each infant the difference in percentage between the looking times to the two objects was also calculated for both conditions. These differences were significantly higher than 0 in both conditions, that is, infants spent more time looking at the object that was acted upon than the other object [Hand: $t(17) = 8.65$, $p < .0001$, two-tailed; Tube: $t(17) = 18.48$, $p < .0001$, two-tailed]. The difference, however, was larger in the Poking Tube condition than in the Poking Hand condition [$F(1, 35) = 9.17$, $p < .005$, effect size: 23.4%]. There was no effect of the Age Group factor. This result indicates that the 'spotlight effect' was in fact stronger in the Poking Tube condition. Therefore, our finding that in the Poking Hand condition infants looked significantly longer in the new goal/old location test event than in the old goal/new location test event cannot be reduced to a 'spotlight effect'.

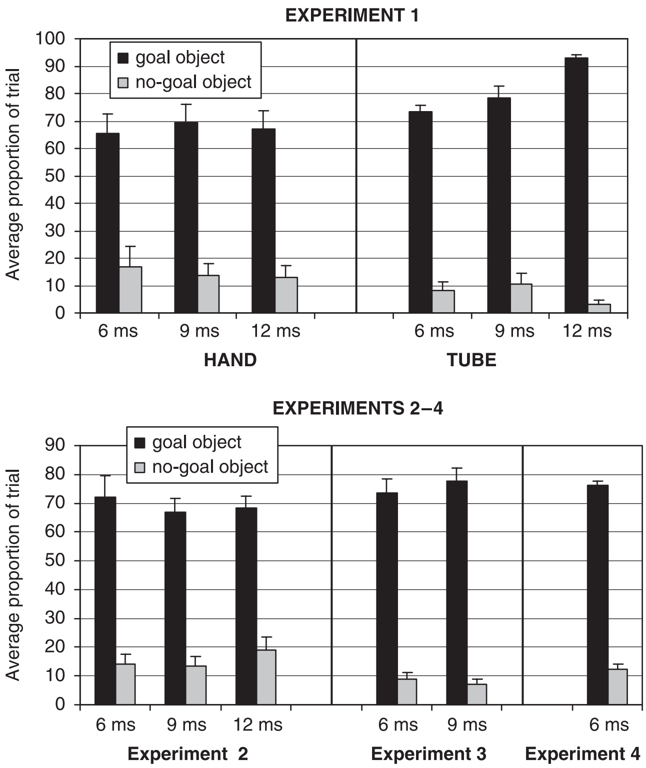


**Figure 3** *Average proportion of test trials infants spent looking at the goal and the no-goal object for each experiment and age group.*

### *Discussion*

In the first experiment we investigated whether infants are able to encode the goal object of an unfamiliar human action (Poking Hand) and an inanimate action (Poking Tube) if these actions provide the infants with a behavioral cue for goal-directedness, namely, equifinal variations of the action towards the goal-state. In the Poking Hand condition, we found that 12-, 9- and even 6-month-old infants can evaluate the goal-directedness of the observed action. The fact that even the youngest age group showed this ability suggests that infants' own experience with particular actions is not a necessary precondition for the ability to interpret human actions as goal-directed. Instead, as we proposed, it is the presence of relevant behavioral cues that allows such an interpretation.

Our result is consistent with the finding of Jovanovic and her colleagues (under review) which showed that infants as young as 6 months old are also able to encode the goal of an unfamiliar hand action when the behavioral cue that was provided was a salient action effect. Thus, there are two independent studies using different behavioral cues providing evidence that it is not only the familiar grasping action that infants at such an early age can consider as goal-directed. Furthermore, the facilitating role of the equifinality cue in interpreting the poking hand action as goal-directed is also supported indirectly by Woodward and Guajardo's recent study (2003). In this study, infants observed a human action in which an extended index finger simply touched the target object once and then rested on the object. It was found that only 12-month-olds and some 9.5-month-olds could consider this action goal-directed. Touching the object with the index finger involves the same movement as our poking action except that it lacks behavioral cues. Perhaps then it was the equifinal variations that were present in our repetitive-from-different-angles poking hand action that triggered goal attribution in infants as young as 6 months old.

In the inanimate Poking Tube condition infants also demonstrated sensitivity to the goal-directedness of the action. This can be concluded from the fact that the looking patterns did not differ across the two conditions (Poking Hand vs. Poking Tube). This finding is an important one, because Woodward's critical result (1998) was an interaction between the inanimate and animate conditions showing that infants treated the two conditions markedly differently in terms of goal-directedness. This was not the case in our present study. On the other hand, when the Poking Tube condition was analyzed separately in each age group, none of the age groups showed evidence for goal-attribution (although the 6- and the 12-month-old infants showed the same pattern in their looking times as in the Poking Hand study, see Figure 2). Therefore, we cannot conclude with confidence that infants demonstrated the ability to interpret the inanimate Poking Tube action as goal-directed.

It is possible that the reason we did not find as strong evidence for goal attribution in the Poking Tube condition as in the Poking Hand condition, is that equifinal variations of the action – the behavioral cue that was provided to help trigger the goal-directed action interpretation – was expressed less effectively in the Poking Tube condition than in the Poking Hand condition. In both conditions, equifinal variations of the action were created by the actor approaching the target toy from various angles and poking the toy on different sides. However, since a human hand is more flexible than a paper tube (because our joints, fingers, hand and arm can all move separately) the variations of the hand's movement might have been more apparent than those of the paper tube. This difference may have resulted in not having the behavioral cue effectively available in the Poking Tube condition.

Another concern regarding the way in which the equifinality cue was applied in our experiment is that it did not include equifinal adjustments. That is, there was no change in the environment or in the properties of the goal object that the actor had to adjust its action to. We simply showed the infants that the actor is capable of achieving its goal via different paths (equifinal variations). Some argue (Gergely *et al.*, 1995; Csibra & Gergely, 1998; Tomasello, 1999; Dennett, 1987) that witnessing equifinal adjustments in a changed situation is essential for being able to predict future goal-directed actions. Therefore, it is conceivable that lacking this characteristic of equifinality made this cue less effective. However, equifinal adjustments were missing in both conditions and yet, infants did attribute goals in the Poking Hand condition.

A further possibility is that in the case of inanimate actions a single cue is not sufficient for eliciting a goal-directed action interpretation. Recall that one group of theories (e.g. Leslie, 1994; Csibra *et al.*, 2003) emphasized that infants often require *multiple* behavioral cues to successfully evaluate goal-directedness of actions. The presence of multiple behavioral cues might be particularly important in the case of inanimate actions, since infants do not have any prior knowledge of inanimate actors, for example, as agents, that is, able to move and change their path autonomously. On the other hand, hands are already categorized as agents before the infant even comes into the laboratory due to the infant's experience with the behavioral characteristics of hands, both his own and other people's. However, such pre-categorization of rods, mechanical claws or paper tubes will not have taken place. Therefore, in the case of inanimate actors, infants may need to be provided with more direct perceptual cues. This possibility is also supported by two 'negative' findings (Jovanovic *et al.*, under review; Kamewari, Kato, Kanda, Ishiguro & Hiraaki, 2005, Experiment 3). In both of these studies inanimate actors displayed only a single cue (an action-effect or self-propulsion, respectively) and 6-month-olds were unable to interpret the actions as goal-directed. In Experiments 2, 3 and 4 we investigated this hypothesis by testing infants' ability to encode the goal object of an inanimate action that displayed the combination of some or all of the following cues: self-propelledness, action-effect and equifinal variations of actions.

## Experiment 2

Woodward's set-up (1998) was applied again. This time we used a wooden rod as the inanimate actor. The action of the rod provided the infants with two of the behavioral cues mentioned above. In the familiarization phase, the rod demonstrated self-propelledness: it freely moved around by changing its path in a random fashion. Then it produced an effect in its goal-approach: it touched, lifted up, and then replaced the target toy. This novel and 'magical' effect was achieved by attaching pieces of Velcro to both the wooden rod and the toys (see Figure 4).

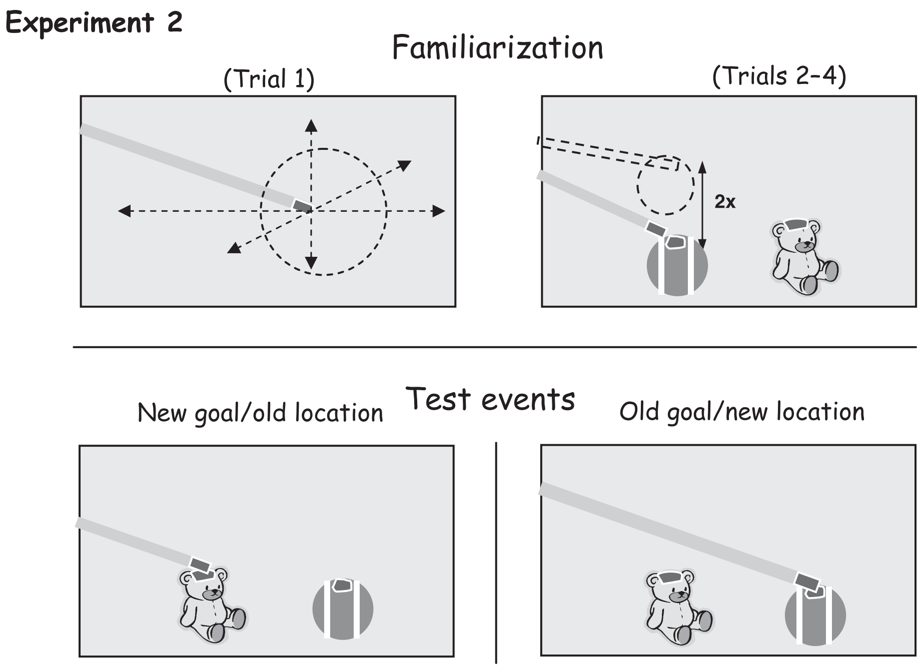


**Figure 4** *Stimulus events in Experiment 2.*

### Method

#### Participants

Thirty-two 12-month-old (19 males and 13 females, mean age = 11 months 26 days, SD = 20.4 days, range 11 months to 12 months and 29 days), 32 9-month-old (17 males and 15 females, mean age = 9 months 2 days, SD = 10.6 days, range from 8 months and 9 days to 9 months and 18 days) and 13 6-month-old (7 males and 6 females, mean age = 6 months 2 days, SD = 14 days, range from 5 months and 17 days to 6 months 28 days) infants participated in this study. An additional ten 12-month-old, eight 9-month-old and one 6-month-old infants were excluded from the data analysis because of fussiness (five, three, zero, respectively), experimental error (eight, four, one), or inattentiveness during familiarization (zero, one, zero).

#### Apparatus

The apparatus used in Experiment 2 was identical to that of Experiment 1, except that the stage was only 55 cm wide and the toys were not sitting on a pedestal but directly on the stage 20 cm apart.

#### Stimuli and procedure

*Familiarization trials.* Infants saw five familiarization trials. The curtain was raised between the trials and there was an approximately 2–3 sec pause between the trials. In the first familiarization trial the stage floor was empty. After a computer signal turned the lights on, the experimenter entered a wooden rod (60 cm long, 2.5 cm wide and 1.5 cm deep) through the opening on the right side of the stage and moved the rod several times in random fashion (up, down and around) in the stage area. The end of the rod (about 7 cm long) was wrapped with red Velcro (see Figure 4). This trial lasted about 8 sec. When the curtain was raised two toys were placed on the stage, a yellow plastic bear and a green ball with white stripes (the same toys as were used in Experiment 1). There was a 3 × 3 cm piece of white Velcro attached on the top of both the bear and the ball. In the second familiarization trial the infants saw the two toys sitting on the stage for 5 sec. In the 3–5 familiarization trials the experimenter reached in with the rod and touched the top of one of the two toys. The rod stayed still for 3 seconds then lifted the toy up (thanks to the Velcro) and held it in the air about 15 cm high for 3 seconds before placing it back on the stage floor and leaving it there for another 3 seconds. The pick-up was repeated a second time. At the end of the trial the rod stayed still in touch with the ball on the stage floor for 5 seconds (see Figure 4). The trial lasted about 18 seconds. The experimenter's hand was not visible in any of the trials. The rod and the toy were separated only when the curtain was raised after the trials. After the fifth familiarization trial the curtain was raised and the positions of the two toys were switched. Then the curtain was lowered and the toys were shown in their new locations but without the presence of the rod for 5 seconds.

*Test trials.* One of two types of test events was given to the infants. Half of the infants saw two identical test trials in which the rod touched the same toy as in the familiarization trial. This time, however, the rod did not lift the toy, but merely stayed still until the infant looked away. However, since toys had been swapped, the touching action took place at a different location (old goal/new location test event). The other half of the infants saw two test trials in which the rod touched the new toy while the action took place at the old location (new goal/old location test event; see Figure 4). The online observer started to measure the looking times when the rod touched the toy and when the infant was looking at the stage. All other aspects of the procedure were identical to those of Experiment 1. To assess reliability, the same procedure was used as in Experiment 1. In all cases the agreement was above 95% either between the online observer and one of the offline observers, or between the two offline observers.

### Results

The looking times of the two test trials were analyzed. Preliminary analysis did not reveal any effect of the location of the poking action (near side vs. far side), the position of the toys (bear left vs. bear right) or the target object (bear vs. ball). Therefore, these factors were omitted from further analyses. Using Trial Type (2) [new goal/old location vs. old goal/new location], Sex (2), Age group (3) [6, 9 and 12 months] as between-subject factors and Test Trial (2) as within-subject factor an ANOVA was carried out. The analysis found a Test Trial main effect [$F(1, 65) = 8.27$, $p < .005$, effect size: 11.3%], indicating that looking times significantly declined across the two test trials. The ANOVA also found an Age group main effect [$F(1, 65) = 3.32$, $p < .05$, effect size: 9.3%]. Post-hoc tests (Bonferroni) revealed that this was due to the fact that overall the 9-month-old infants looked significantly longer than the 12-month-olds. Furthermore, there was a Trial Type × Sex interaction [$F(1, 65) = 8.40$, $p < .005$, effect size: 11.4%], indicating that overall girls looked in the new goal/old location test event longer than in the old goal/new location test event

[$F(1, 32) = 5.96$, $p < .02$, effect size: 15.7%], while boys looked equally long in the two type of test events.[1] Finally, the ANOVA found a significant Test Trial × Trial Type × Age group interaction [$F(1, 65) = 3.37$, $p < .04$, effect size: 9.4%] as well as a Trial Type × Age group × Sex interaction [$F(1, 65) = 3.31$, $p < .05$, effect size: 9.3%].

To explore these interactions, separate ANOVAs were carried out in each age group with the following results: In the 12-month-old age group infants looked significantly longer in the new goal/old location test trial than in the old goal/new location test trial [$F(1, 28) = 11.63$, $p < .002$, effect size: 29.3%] in the first test trial. The 9-month-old age group showed a similar pattern: they also looked longer in the new goal test event that in the old goal test event, but the difference fell short of significance [$F(1, 28) = 3.55$, $p < .07$, effect size: 11.3%]. However, non-parametric tests revealed a reliable difference between the new goal/old location test event and the old goal/new location test event in the first test trial in this age group (Mann-Whitney U = 72.0, $z = -2.11$, $p < .04$, two-tailed).[2] The 6-month-old age group looked equally long in the two types of test events. Sex had no effect on the looking times when the age groups were analyzed separately, except in the 9-month-old age group where a tendency for a Trial type × Sex interaction was found in the first trial [$F(1, 28) = 3.36$, $p < .077$, effect size: 10.07%]. Mean looking times with standard errors by age groups in the first test trials are shown in Figure 5.

As in Experiment 1, the 'spotlight effect' of the rod was explored in Experiment 2. Eighteen infants (six from each age group) were randomly selected and their video recordings were scored in the same way as in Experiment 1. Mean proportions of the total trial time spent looking at the object that was touched by the rod and at the other object are depicted on Figure 3. The difference between the proportional looking times to the two objects was significantly higher than 0, that is, infants spent more time looking at the object that was acted upon than the other object [$t(17) = 13.00$, $p < .0001$, two-tailed]. There was, however, no difference between the three age groups [$F(1, 17) = 0.34$, $p < .71$, effect size: 4.4%]. These results indicate that the rod drew infants' attention to the target object and that it served as a spotlight with the same effectiveness in all age groups. Note, however, that only the 12-month-old age group looked significantly longer in the new goal/old location test event than in the old goal/new location test event. If 'spotlight effect' were responsible for this looking pattern, then the other two age groups should also have looked longer at the new goal/old location test event. Since this was not the case, the 'spotlight effect' is an unlikely explanation for the main finding in Experiment 2.

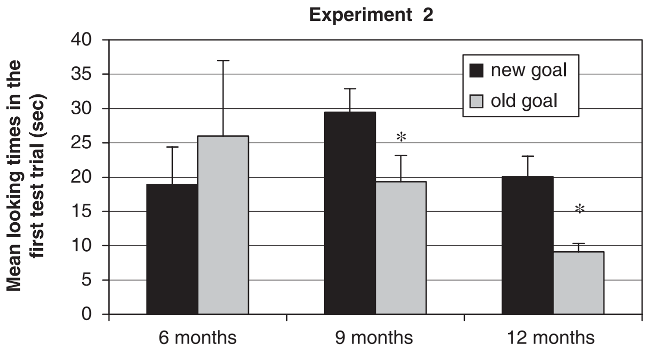


**Figure 5** *Mean looking times (bars show SEM) in the first test trial in Experiment 2 as a function of Trial Type and Age group (* = p < .05, only with non-parametric test in 9-month-olds).*

### *Discussion*

In Experiment 2 we tested whether infants can attribute a goal to an inanimate action that exhibits two behavioral cues: self-propelledness and an action-effect. We found that 12-month-old infants were able to encode the goal of the inanimate action when these cues were provided. Nine-month-olds also showed a similar ability, as revealed by non-parametric test. Six-month-olds, however, were not able to evaluate the goal-directedness of the inanimate action. The performance of the 12-month-old infants in Experiment 2 compared to that in the Poking Tube condition of Experiment 1 suggests that the presence of multiple cues as opposed to a single cue results in an enhanced ability to identify the goal of an inanimate action. Why did the 6-month-olds not benefit from the availability of multiple cues in Experiment 2? We argue that their difficulty might be due to certain aspects of our experimental design rather than to their insensitivity to these cues.

Recall that the two cues were available only during the familiarization period. In the test events the wooden rod simply touched the toy and stayed still. Thus, there were no action-effect (pick-up) or self-propelledness cues provided in the test events. One can speculate that when the infants watched the test events, they expected the rod to

[1] There are only a few studies reporting sex differences in various cognitive or perceptual tasks in infancy (e.g. Baillargeon & DeVos, 1991; Kavsek, 2004; Arterberry & Bornstein, 2002). There are no clear explanations for these differences; neuropsychological reasons are sometimes suggested with certain brain structures being different due to hormones. Since sex difference was only found in Experiment 2 and not in any of the other three experiments, we cannot make a general conclusion of girls being better/quicker than boys in their development of goal-attribution.

[2] In all experiments, non-parametric tests were also carried out to confirm the results. These tests are only reported when they are different from the results of the parametric tests.

act the same way as in the familiarization, that is, they expected it to pick up a toy (regardless of which toy the rod approached). The fact that the rod's action differed from their expectations may have diverted the infants' attention from identifying which particular toy was involved. Thus, their looking times in the test events may have reflected the change in the action of the rod, which, however, was the same in both types of test events (new goal vs. old goal). Apparently, the 12-month-old infants were able to overcome their expectations and therefore paid attention to the particular toy involved in the rod's action.

Another possible explanation for our finding is related to infants' ability to interpret actions as sequences in the service of an overarching goal. This ability was elegantly demonstrated by Woodward and Somerville (2000), who showed that an action – which by itself was *not* treated as goal-directed by infants – can be considered as goal-directed by 12-month-olds if they have already seen the action in a goal-directed context (if it was a prerequisite for achieving a goal, a means to an end). In our experiment, the 'rod touching the toy' can be considered as a sub-action in the service of the goal action: the 'rod picking up the toy'. It is then conceivable that the 12-month-old infants could link the sub-action to the goal action in the familiarization. Then in the test events, without witnessing the goal action itself, they were able to interpret the sub-action as directed to the same goal as before, and consequently they were surprised to see the different target object. On the other hand, the 9-, and particularly the 6-month-old infants could not relate the sub-action as a causal prerequisite for the goal and seeing a different toy being acted upon did thus not violate their expectations.

## Experiment 3

With these possible explanations for the results of Experiment 2 in mind, Experiment 3 replicated Experiment 2 except that the rod not only touched but also picked up the toy in the test events. That is, the action-effect cue was available in both the familiarization phase and in the test events. This way we could exclude the possibility that the younger infants were unable to encode the goal object of the inanimate action because of difficulties in linking the sub-action to the goal action. We tested 9- and 6-month-old infants.

### *Method*

#### Participants

Twenty-four 9-month-old (15 males and 9 females, mean age = 8 months 30 days, SD = 10.8 days, range 8 months 16 days to 9 months and 27 days) and 13 6-month-old (6 males and 7 females, mean age = 6 months and 1 day, SD = 11.1 days, range from 5 months and 16 days to 6 months and 15 days) infants participated in the study. An additional six 9-month-old and two 6-month-old infants were also tested but were excluded from the data analysis because of fussiness (one, zero, respectively) or experimental error (five, two).

#### Apparatus

The apparatus used in Experiment 3 was identical to that of Experiment 2.

#### Stimuli and procedure

The stimuli in Experiment 3 were identical to those of Experiment 2 with the following exception. In both types of test events the rod did not only touch the target toy but also lifted it up once before placing it back in the same manner as in the familiarization trials (see Figure 6). The online observer started to measure the looking times when the rod touched the toy and when the infant was looking at the stage. To assess reliability, the same procedure was used as in Experiment 1. In all cases the agreement was above 95% either between the online observer and one of the offline observers, or between the two offline observers.

### *Results*

The looking times of the two test trials were analyzed. Preliminary analysis did not reveal any effect of sex, location of the poking action (near side vs. far side), the position of the toys (bear left vs. bear right) and the target object (bear vs. ball). These factors were therefore omitted from further analyses. Using Trial Type (2) [new goal/old location vs. old goal/new location] and Age group (2) [9- vs. 6-month-olds] as between-subject factors and Test Trial (2) as within-subject factor, an ANOVA was carried out. It revealed a Test Trial main effect [$F(1, 21) = 4.64$, $p < .04$, effect size: 18.1%], indicating that overall infants looked longer in the first test trial than in the second test trial. A Trial Type × Age group interaction was also found [$F(1, 21) = 4.35$, $p < .05$, effect size: 17.2%]. This interaction was further investigated by separate ANOVAs in each age group. In the 6-month-old age group, no main effects or interactions were found. In the 9-month-old age group the Trial Type showed a significant main effect [$F(1, 22) = 4.29$, $p < .05$, effect size: 16.3%], indicating that 9-month-old infants looked longer in the new goal/old location test event than in the old goal/new location test event. The mean looking times

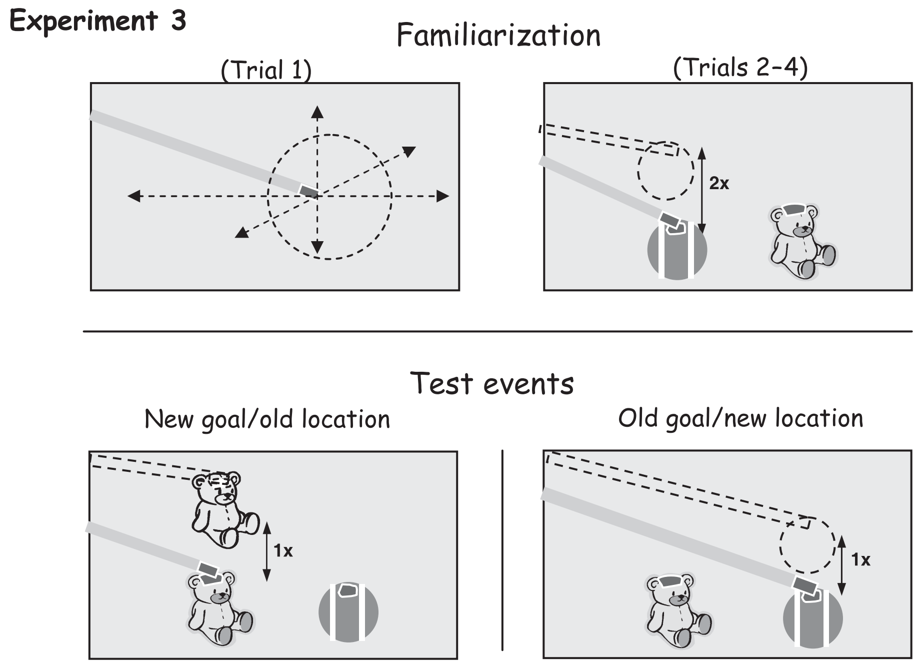


**Figure 6** *Stimulus events in Experiment 3.*

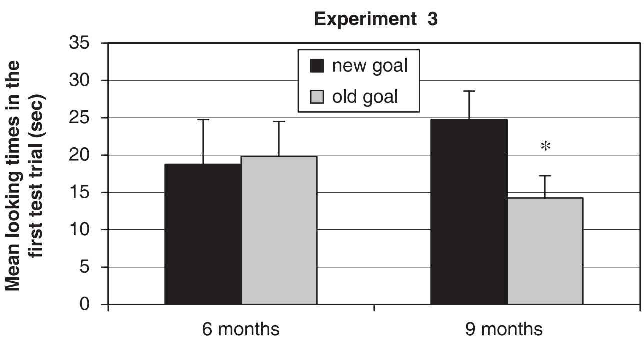


**Figure 7** *Mean looking times (bars show SEM) in the first test trial in Experiment 3 as a function of Trial Type and Age group (* = p < .05).*

with standard errors in the first test trials are shown by age groups in Figure 7.

As in the previous experiment, the 'spotlight effect' of the rod was also explored in Experiment 3. Eighteen infants (nine from each age group) were randomly selected and their video recordings were scored in the same way as in Experiment 1. Mean proportions of the total trial time spent looking at the object that was lifted by the rod and at the other object are depicted in Figure 3. The difference between the proportional looking times to the two objects was significantly higher than 0, that is, infants spent more time looking at the object that was acted upon than the other object [$t(17) = 15.58$, $p < .0001$]. The differences between these looking times, however, did not differ in the two age groups [$F(1, 17) = 0.43$, $p < .52$, effect size: 2.6%]. These results indicate that the rod's action was just as effective a spotlight for the 6-month-old age group as it was for the 9-month-old age group. However, it was only the 9-month-old group that looked significantly longer in the new goal/old location test event than in the old goal/new location test event. The same argument can be made here as in Experiment 2: looking patterns for the new goal/old location and the old goal/new location test events cannot be caused by the spotlight effect, because then 6-month-old infants should have also looked longer in the new goal/old location test event than in the old goal/new location test event.

### *Discussion*

In Experiment 3 we found that making the action-effect cue available not only in the familiarization but also in the test events apparently enhanced the size of the effect in 9-month-olds to attribute goals to the inanimate action. In Experiment 3, the effect size increased from 11% to 16% and was now significantly different between the looking times for the two test events on parametric tests. This indicates that when the rod not only touched but also picked up the toy in the test event, the infants expected the rod's action to be directed to the same target object as in the familiarization. This suggests that 9-month-olds still benefit from direct perceptual evidence for the entire goal-directed action sequence in identifying the goal.

The 6-month-olds, however, did not seem to be helped by the addition of the action-effect cue to the test events. Their looking times did not reflect that they considered the rod's action as goal-directed. Are infants not able to encode the goal of an inanimate action at all at such an early age or were they still lacking the appropriate stimuli? To probe this further, our next step was to enrich the inanimate action with additional behavioral cues to see whether the latter explanation was correct. In Experiment 1, the equifinality cue was successfully used by 6-month-olds in the case of the animate 'poking hand' action and, even in the inanimate 'poking tube' condition, 6-month-olds showed a tendency to be sensitive to this cue. Therefore, in Experiment 4 we tested the power of having available all three of the cues that we used so far: self-propelledness, action-effect and equifinality.

## Experiment 4

We replicated Experiment 3 with the exception that a third cue, equifinal variations of the action, was added to the familiarization phase. The wooden rod approached the target toy from three different angles and picked the toy up by sticking to three different parts. We expected that having all three cues available, 6-month-old infants might be able to consider the inanimate action as goal-directed.

### *Method*

#### Participants

Twenty-two 6-month-old (15 males and 7 females, mean age = 6 months 5 days, SD = 17.7 days, range from 5 months and 16 days to 7 months and 2 days) infants participated in the study. One additional infant was also tested but was excluded from the data analysis because of inattentiveness during familiarization.

#### Apparatus

The apparatus used in Experiment 4 was identical to that of Experiments 2 and 3.

#### Stimuli and procedure

The stimuli in Experiment 4 were identical to those of Experiment 3 with the following exception. There were Velcro pieces attached to the top, the front and the back of both toys. In the familiarization trials the rod entered the stage area from various angles and lifted the target toy by touching either its top (fam. trial 3), front (fam. trial 4) or back (fam. trial 5; see Figure 8). The test trials were identical to those of Experiment 3. To assess reliability, the same procedure was used as in Experiment 1. In all cases the agreement was above 95% either between the online observer and one of the offline observers, or between the two offline observers.

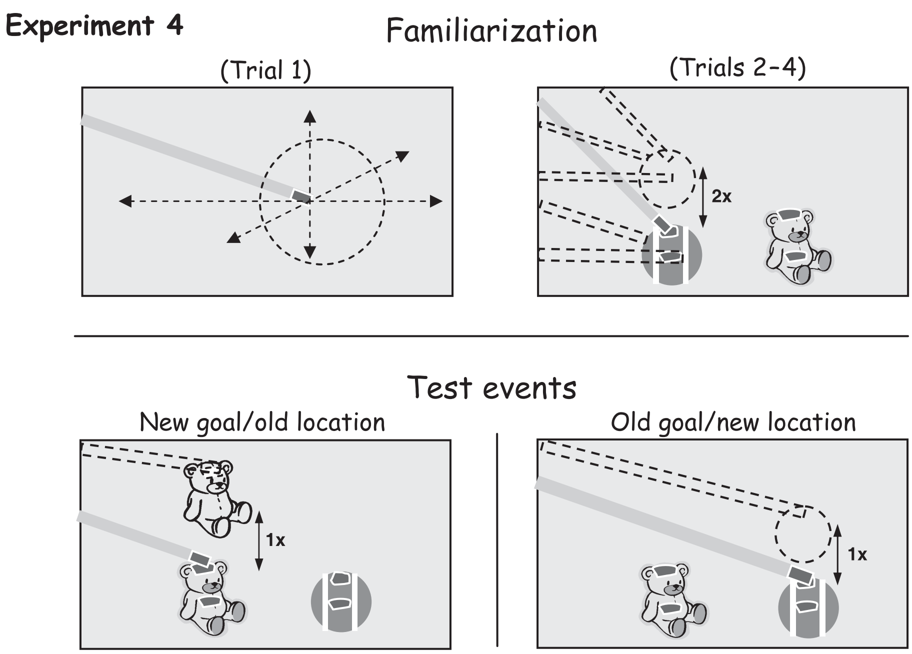


**Figure 8** *Stimulus events in Experiment 4.*

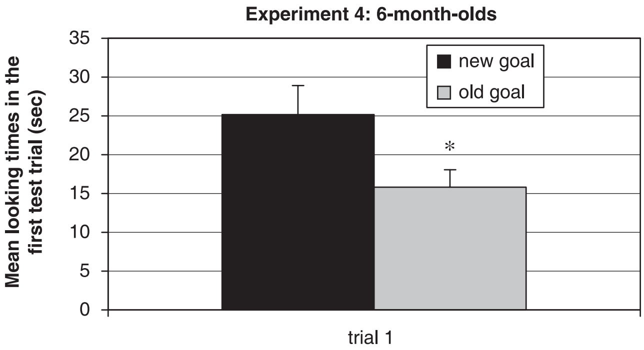


**Figure 9** *Mean looking times (bars show SEM) in the first test trial in Experiment 4 as a function of Trial Type (* = p < .05).*

### *Results*

The looking times of the two test trials were analyzed. Preliminary analysis did not reveal any effect of sex, location of the poking action (near side vs. far side), the position of the toys (bear left vs. bear right) and target object (bear vs. ball). Therefore, these factors were omitted from further analyses. Using Trial Type (2) [new goal/old location vs. old goal/new location] as between-subject factor and Test Trial (2) as within-subject factor an ANOVA was carried out. A Trial Type main effect was found [$F(1, 20) = 5.18$, $p < .034$, effect size: 20.6%], indicating that infants looked longer in the new goal/old location test event than in the old goal/new location test event. The mean looking times with standard errors in the first test trial are shown in Figure 9.

Since there was only one age group in this experiment, the spotlight effect of the rod was compared with that of the previous experiments. Eighteen infants were randomly selected from Experiment 4 and their video recordings were scored in the same way as in the previous experiments. Mean proportions of the total trial time spent looking at the object that was lifted by the rod and at the other object are depicted in Figure 3. The difference between the proportional looking times to the two objects was significantly higher than 0, that is, infants spent more time looking at the object that was acted upon [$t(17) = 14.73$, $p < .0001$, two-tailed]. The scores of the 6-month-old age groups across the four experiments were compared using Scheffe post-hoc test which showed that these were not different [$F(1, 44) < 1$, *ns*]. This indicates that the same size of spotlight effect was elicited in 6-month-olds in all experiments. Thus, the fact that the four experiments involved different actions (presenting different behavioral cues) did not affect the strength of the spotlight effect in 6-month-olds. To test if goal-directed action interpretations of the actions were influenced by the size of the spotlight effect, those groups in which 6-month-old infants looked significantly longer in the new goal/old location test event than in the old goal/new location test event (Experiments 1 and 4) were compared to those where no significant difference was found between the two test events (Experiments 2 and 3). The ANOVA revealed no difference in spotlight effect between these groups, indicating that the relative novelty of the new object was not responsible for the looking patterns for the test events [$F(1, 44) < 1$, *ns*].

Across experiment comparisons were also carried out to see if the type of action had any impact on the size of spotlight effect. The ANOVA comparing the four experiments revealed no difference in the spotlight effect (in the scores for proportional looking time differences for the two objects) between the experiments [$F(1, 89) = 1.53$, $p < .21$, effect size = 5.1%]. To see whether the spotlight effect had any role in the particular looking patterns of infants for the two test events we compared the size of the spotlight effect for those groups that showed a significant difference between the two test events across the four experiments and those that did not. The ANOVA showed that there was no difference in the size of spotlight effect between these groups [$F(1, 89) = 2.38$, $p < .12$, effect size = 2.6%]. Taken together, all the analyses in the four experiments suggest that the spotlight effect can be excluded as an alternative explanation for the pattern of looking times for the test events.

### *Discussion*

The results of Experiment 4 suggest that even very young infants are capable of interpreting inanimate actions as goal-directed. However, it seems that to do so they require the presence of multiple cues. While equifinality by itself (see Experiment 1, Poking Tube condition), or self-propelledness coupled with an action-effect (Experiments 2 and 3) were not sufficient to elicit goal-directed action interpretation from 6-month-old infants, the simultaneous presence of these three cues was. The question of how and to what extent the ability of goal attribution depends upon these particular cues will be considered in the general discussion.

## General discussion

In four experiments we investigated the nature of goal-directed action interpretation in infancy. In particular, the experience-based and the cue-based accounts were contrasted by examining certain action characteristics that are assumed to be critical in goal-attribution by one or the other account. Three main questions were raised.

One concerned the familiarity of the action. We found that infants as young as 6 months old can consider an action goal-directed even if the action is not likely to occur in their everyday life and if they cannot perform it themselves. This finding does not comport with the experience-based approach: infants do not necessarily have to acquire prior, hands-on experience with particular actions in order to be able to encode the goal of the action. The second question concerned the perceptual appearance (animate vs. inanimate) of the actor. We found that (under certain circumstances, see below) 12-, 9- and even 6-month-old infants can consider the action of both types of actors as goal-directed.

Our finding that infants are able to judge the goal-directedness of the action of an inanimate object as early as 6 months is also supported by some recent studies (Luo & Baillargeon, 2005; Kamewari *et al.*, 2005, Experiment 2). These findings suggest that ascribing goal-directedness is initially independent from the specific appearance of the actor and not restricted to human actions. In an important demonstration, Luo and Baillargeon (2005), using the Woodward paradigm, showed that 5-month-old babies attributed goals to an object whose motions appeared to be self-propelled. The object was a rigid box without face or other animate-like markings that moved back and forth without visible means of propulsion. Self-propulsion was thus the only cue to agency and this proved sufficient for young babies to attribute goals. Note that 5 months is the earliest age for which sensitivity to the goal-directed nature of inanimate actions has been reported. Evidence for such an early onset is important for two reasons. First, it confirms the hypothesis that previous studies which found a lack of goal attribution in case of inanimate actions before 9 months of age reflected difficulties with task demands rather than a failure to encode the goal of the action by relying on abstract behavioral cues (e.g. Csibra *et al.*, 1999). Second, since the earliest evidence for goal attribution in the case of human actions was also found at around 5 or 6 months (except from one study, but see argument below), the assumption of the cue-based account, namely that infants apply goal-directed interpretation to actions of a wide range of entities *from the start*, is further supported.

The third question concerned the influence of behavioral cues on infants' ability to infer the goal of an action. Three cues were used that have been widely suggested in the literature as critical to goal-attribution: self-propelledness, equifinal variations of the action, and a salient action-effect. Overall, our findings demonstrate that the presence of these cues can elicit infants' goal-directed action interpretations, while the lack of these cues can prevent them from interpreting actions as goal-directed (see below for specification). The performance of the infants across the four experiments also provides support for the view that it is not a single cue that automatically triggers goal attribution, but rather that infants use multiple cues when they make a judgment about goal-directedness. This was demonstrated clearly in Experiment 4, where we found that the youngest age group, the 6-month-olds, could only consider an inanimate action as goal-directed when all three cues were present.[3]

The relationship between the cues and the role they play in determining goal-directedness is currently under debate. Particularly, the presence of self-propelledness was proposed by some (Premack, 1990; Baron-Cohen, 1994) as a precondition for setting up a goal-directed interpretation: only when an object's motion is self-propelled will infants consider other cues that are relevant for identifying the goal of the action. Others (Csibra *et al.*, 1999), however, demonstrated that for 9- and 12-month-old infants, witnessing an object's movement as self-initiated is neither a sufficient nor a necessary condition for making inferences about the object's goal-directed behavior (see Luo & Baillargeon, 2005). Note that because our present study did not investigate the relationship (necessary and sufficient conditions) between the three cues systematically, we cannot draw strong conclusions about this issue. The findings in Experiment 1 were inconclusive regarding the sufficiency of the equifinality cue by itself. The self-propulsion cue and the action-effect cue were not available solo in any of the experiments. The simultaneous presence of these two cues was sufficient for 9- and 12-month-old infants to encode the goal of the action (see Experiments 2 and 3). Finally, Experiment 4 showed that when all three cues were available, even 6-month-olds did interpret the action as goal-directed. (We will return to the possible reasons for the age-related differences.) However, whether all three cues are necessary, or whether either 'equifinality with self-propulsion' or 'equifinality with an action-effect' is sufficient for goal-attribution in 6-month-olds is an open question. In fact, Kiraly and her colleagues (2003) suggested that the latter pair, 'effect plus equifinality', is sufficient, which is consistent with our finding. Overall, equifinal variations may be relatively more important for goal attribution than the other cues in our study. One may speculate that what equifinal variations convey is an impression of 'autonomous behavior' or 'free will'. In Shimizu and Johnson (2004) infant goal attribution

[3] Note that in Experiment 1, the equifinality cue by itself was sufficient to encode the goal of the Poking hand action. However, as we argued in the introduction, the infants might also have relied on other cues already associated with the actions of human hands. Therefore, the multiplicity question can be safely addressed only in the case of unfamiliar inanimate actions.

appeared to follow cues that an unfamiliar object made a 'choice' between two targets. Self-initiated movement too may be effective just insofar as it conveys this 'autonomy'. Studies with adults have begun to characterize the psycho-physical basis of what we are calling 'cues' (e.g. Tremoulet & Feldman, 2000). Studies with infants are needed to determine their developmental role.

Our findings also accord with the idea that the relationship between cues and the identification of goal-directedness is probabilistic (Leslie, 1995; Csibra, 2003). While the presence of some cues might indeed be necessary, other cues may simply make it more (or less) likely that infants will ascribe goal-directedness to an action. That is, the presence of additional cues can strengthen the interpretation of the action as goal-directed. The improvement in the performance of our youngest infants as more cues became available (Experiments 2 to 4) may reflect just such a relationship.

We turn now to the function of behavioral cues in the development of goal attribution. Adults' goal-directed action interpretations do not (entirely) depend on the direct perception of behavioral cues (although in ambiguous cases we do look for them) and is (mostly) applied to human action. We argue that this shift takes place gradually as infants start to calculate the statistical association between the cues they are initially sensitive to and between these and other event properties, such as the type of actors that are likely to exhibit these cues. When such associations are established, infants can anticipate goal-directed actions from actors identified by surface features without collecting direct evidence for behavioral cues. We propose then that the perception of goal-directed action develops by bootstrapping from an innate cue base into a learning mode. In this view, the infant begins with a core notion of goal-directed agent that is triggered by a certain range of cues. This provides the entry point into a domain-specific learning process that accretes further information about the different kinds of agents and actions the infant is exposed to. Furthermore, such learning, expanding outward from the innate base, can go forward efficiently: featural cues specifying a hand can come to signal goal-directed agency via association with the innate behavioral cues. An infant, who has already categorized hands as agents, can then anticipate that a hand will *tend* to act in a goal-directed way, even in the absence of the core behavioral cues. Of course, the mere fact that an infant has categorized a hand as agent in no way guarantees that she will be able to determine for each and every hand motion what its goal might be. New actions will demand new goal learning. One important heuristic for this type of learning might be to hypothesize that the observed typical *effect* of a given action performed by a hand-agent is its intended goal.

As an example of cue-based bootstrapping, consider infants' success with encoding the goal of the *stationary* grasping hand in Woodward's (1998) study. Although there were no (or insufficient) motion cues available in that event considered in isolation, infants had already learned additional cues associated with the hand and with its grasping action. Similarly, the developmental trend found in Experiments 2 and 3 shows that older infants do not require the presence of direct behavioral cues once they have had the opportunity to associate these with the actor – even with an inanimate actor. Recall that 12-month-old infants, but not 6-month-olds, could interpret the stationary touch of the wooden rod as goal-directed after they had seen the wooden rod being able to pick up the target toy, while 9-month-olds appeared to be in transition. That is, 12-month-olds quickly learnt that the wooden rod can produce an effect and could use this knowledge in their reasoning about the goal-directedness of the rod's action even when the effect was no longer present.

Finally, a similar associative learning might have taken place between the appearance of a novel object and its effect in Sommerville and colleagues' study (2005) mentioned before. Recall that in this study 3-month-olds (who typically cannot produce successful goal-directed reaches) could encode the goal of a grasping hand that was wearing a mitten. However, they could only do this if before the test session they had themselves manipulated objects wearing the same mitten that can lift up objects by sticking to them. The authors also offered an alternative to the facilitating effect of infants' own experience to explain their findings. Infants might have learnt the causal characteristics of the mitten, that is, that it can produce these effects on the environment, and then were able to rely on that when they observed a grasping hand wearing the same mitten even without witnessing the effect. The use of tools extends the goal-directed activities of human agents. By 10 months, infants are able to infer an unseen hand as the cause of the motion of an inanimate object (Saxe, Tenenbaum & Carey, 2005), suggesting that infants are equipped early to understand even incompletely observed tool use. Our current work focuses on determining the ways in which such learning can be achieved. By uniting an innate base with this learning process, cue-based bootstrapping can help reconcile divergent views on the emergence of infants' ability to understand actions, even those extended through rods and other tools, as goal-directed.

## Acknowledgment

This work was supported by NSF grant BCS 0079917 awarded to AML.

## References


Arterberry, M.E., & Bornstein, M.H. (2002). Variability and its sources in infant categorization. *Infant Behaviour and Development*, **25**, 515–528.

Baillargeon, R., & DeVos, J. (1991). Object permanence in young infants: further evidence. *Child Development*, **62**, 1227–1246.

Baron-Cohen, S. (1994). How to build a baby that can read minds: cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive/ Current Psychology of Cognition*, **13**, 1–40.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy*. New York: Academic Press.

Berry, D.S., & Springer, K. (1993). Structure, motion, and preschoolers' perception of social causality. *Ecological Psychology*, **5**, 273–283.

Blake, J., O'Rourke, P., & Borzellino, G. (1994). Form and function in the development of pointing and reaching gestures. *Infant Behavior and Development*, **17**, 195–203.

Butterworth, G., Franco, F., McKenzie, B., Graupner, L., & Todd, B. (2002). Dynamic aspects of visual event perception and the production of pointing by human infants. *British Journal of Developmental Psychology*, **20**, 1–24.

Castelli, F., Frith, C., Happe, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, **125** (8), 1839–1849.

Csibra, G. (2003). Teleological and referential understanding of action in infancy. *Philosophical Transactions of the Royal Society (London) B*, **358**, 447–458.

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

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

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

Dennett, D. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

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

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

Guajardo, J.J., & Woodward, A.L. (2004). Is agency skin deep? Surface attributes influence infants' sensitivity to goal-directed action. *Infancy*, **6** (3), 361–384.

Hannan, T.E. (1987). A cross-sequential assessment of the occurrences of pointing in 3- to 12-month-old human infants. *Infant Behavior and Development*, **10**, 11–22.

Heider, F. (1958). *The psychology of interpersonal relations*. Oxford: John Wiley.

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

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

Johnson, S.C., Booth, A., & O'Hearn, K. (2001). Inferring the goals of a nonhuman agent. *Cognitive Development*, **16**, 637–656.

Johnson, S., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science*, **1** (2), 233–238.

Jovanovic, B., Király, I., Elsner, B., Gergely, G., Prinz, W., & Aschersleben, G. (under review). The role of effects for infants' perception of action goals.

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

Kavsek, M. (2004). Infant perception of object unity in static displays. *International Journal of Behavioral Development*, **28**, 538–545.

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.

Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science*, **14** (5), 402–408.

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

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

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

Leslie, A.M. (1995). A theory of agency. In D. Sperber, D. Premack & A.J. Premack (Eds.), *Causal cognition: A multi-disciplinary debate* (pp. 121–141). Oxford: Clarendon Press.

Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science*, **16** (8), 601–608.

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

Meltzoff, A.N., & Moore, M. (1994). Imitation, memory, and the representation of persons. *Infant Behavior and Development*, **17** (1), 83–99.

Michotte, A. (1946/English translation). *The perception of causality*. New York: Basic Books.

Montgomery, D.E., & Montgomery D. (1999). The influence of movement and outcome on young children's attributions of intention. *British Journal of Developmental Psychology*, **17** (2), 245–261.

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*, **36**, 1–16.

Rochat, P., Morgan, R., & Carpenter, M. (1997). Young infants' sensitivity to movement information specifying social causality. *Cognitive Development*, **12**, 441–465.

Saxe, R., Tenenbaum, J.B., & Carey, S. (2005). Secret agents: inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science*, **16**, 995–1001.

Schlottman, A., & Surian, L. (1999). Do 9-month-olds perceive causation-at-a-distance? *Perception*, **28**, 1105–1114.

Scholl, B., & Tremoulet, P.D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, **4** (8), 299–309.

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

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

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

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

Tremoulet, P.D., & Feldman, J. (2000). Perception of animacy from the motion of a single object. *Perception*, **29** (8), 943–951.

Trevarthen, C. (1977). Descriptive analysis of infant communicative behavior. In H.R. Schaffer (Ed.), *Studies in mother–infant interaction* (pp. 227–270). London: Academic Press.

Von Hofsten, C. (1983). Catching skills in infancy. *Journal of Experimental Psychology: Human Perception and Performance*, **9**, 75–85.

Von Hofsten, C. (1991). Structuring of early reaching movements: a longitudinal study. *Journal of Motor Behavior*, **23**, 280–292.

Von Hofsten, C., & Fazel-Zandy, S. (1984). Development of visually guided hand orientation in reaching. *Journal of Experimental Child Psychology*, **38**, 208–219.

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

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

Woodward, A.L., & Guajardo, J.J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development*, **17**, 1061–1084.

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

Woodward, A., Sommerville, J.A., & Guajardo, J.J. (2001). How infants make sense of intentional action. In B.F. Malle, L.J. Moses, & D.A. Baldwin (Eds.), *Intentionality: A key to human understanding* (pp. 149–169). Cambridge, MA: MIT Press.

Received: 24 March 2005
Accepted: 26 April 2006