Cognition 160 (2017) 35–42

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Original Articles

Six-month-old infants expect agents to minimize the cost of their actions

Shari Liu [*], Elizabeth S. Spelke

Department of Psychology, Harvard University, United States

ARTICLE INFO

Article history:
Received 22 May 2016
Revised 9 December 2016
Accepted 20 December 2016

Keywords:
Cognitive development
Goal inference
Social cognition
Open data
Open materials

ABSTRACT

Substantial evidence indicates that infants expect agents to move directly to their goals when no obstacles block their paths, but the representations that articulate this expectation and its robustness have not been characterized. Across three experiments (total $N = 60$), 6-month-old infants responded to a novel, curvilinear action trajectory on the basis of its efficiency, in accord with the expectation that an agent will move to its goal on the least costly path that the environment affords. Infants expected minimally costly action when presented with a novel constraint, and extended this expectation to agents who had previously acted inefficiently. Infants' understanding of goal-directed action cannot be explained alone by sensitivity to specific features of agent's actions (e.g. agents tend to move on straight paths, along supporting surfaces, when facing their goals directly) or extrapolations of agents' past actions to their future ones (e.g. if an agent took the shortest path to an object in the past, it will continue to do so in the future). Instead, infants' reasoning about efficiency accords with the overhypothesis that agents minimize the cost of their actions.

1. Introduction

Action understanding is a fundamentally underdetermined problem: an infinite combination of causes could explain a given observed behavior, including the emotions, desires, and beliefs internal to agents, the goals and obstacles in the world, the physical forces that agents must overcome to achieve their goals, and the forces that their actions produce. In spite of this computational challenge, we solve this problem quickly and intuitively every day: Viewing a simple behavior, like a person walking into a building, licenses inferences about her desires to reach her destination, beliefs about what is there, and competence in planning this action. The building blocks of these capacities emerge early in human development: Infants interpret agents' actions by leveraging assumptions about their material properties (e.g. agents are solid and thus face physical constraints; Saxe, Tzelnic, & Carey, 2006), their causal powers (e.g., agents bring about changes in the motions and states of objects; Muentener & Carey, 2010; Saxe, Tenenbaum, & Carey, 2005) and their goals (e.g. agents face, perceive, and act on objects; Csibra & Volein, 2008; Gergely, Nádasdy, Csibra, & Bíró, 1995; Luo & Johnson, 2009; Woodward, 1998).

These findings raise two important questions about the cognitive infrastructure supporting early action understanding. First, what representations express infants' assumptions about agents and their actions? That is, what are the variables and functions that embody their knowledge? Second, is this content embedded in a coherent system of knowledge or does it reflect local learning about specific actions or physical contexts? In other words, to what extent does this knowledge capture the hidden causal structure of the world versus the statistical regularities in the immediately perceivable environment? The answers to these questions bear on theories of the form and content of mature intuitive psychology, as well as theories of its development.

1.1. Case study: Unpacking rational agency

The assumption that agents seek to maximize rewards and minimize costs, given their beliefs about state of the world, has long been proposed as a key principle in intuitive psychology (Baker, Saxe, & Tenenbaum, 2009; Dennett, 1987; Gergely & Csibra, 2003; Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016). By 5 years of age, we attribute mental states, beliefs and desires, to agents (Bartsch & Wellman, 1989; Wellman, Cross, & Watson, 2001) by assuming that they have planned their actions so as to bring about maximum utility (Jara-Ettinger, Gweon, Tenenbaum, & Schulz, 2015). Nevertheless, questions about the origins of this capacity remain open. Does infants' earliest understanding of agents center on the assumption that their actions are guided by

---

[*] Corresponding author at: William James Hall, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, United States.

E-mail address: shariliu01@g.harvard.edu (S. Liu).

http://dx.doi.org/10.1016/j.cognition.2016.12.007

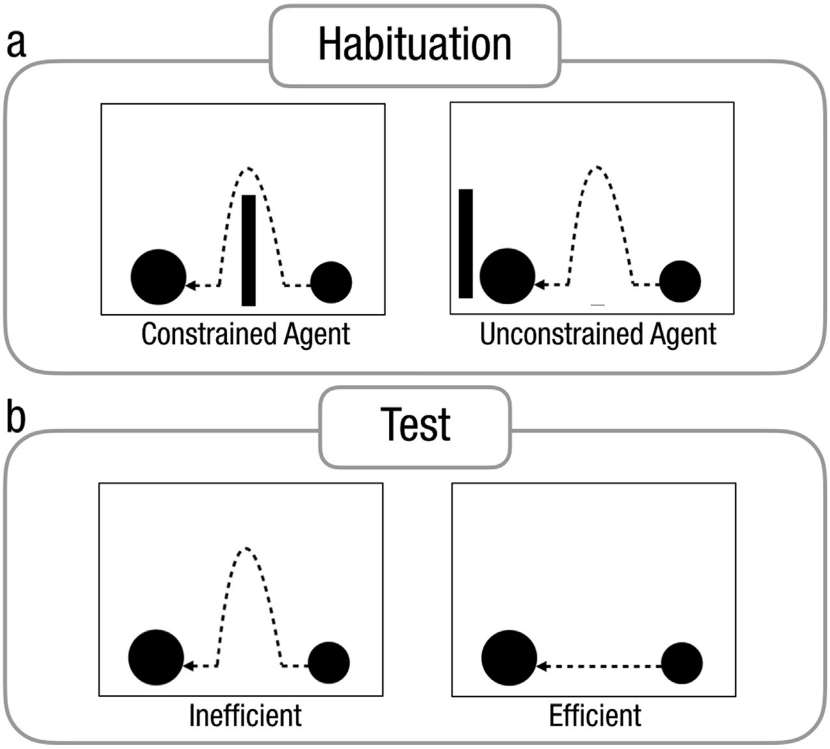

Fig. 1. Schematic of events used in past experiments (e.g. Csibra et al., 1999; Gergely et al., 1995) probing infants' sensitivity to action efficiency. Infants were (a) habituated to an agent leaping over an obstacle (left), or to an agent performing the same actions with the obstacle situated beyond its goal (right), and (b) then viewed test events where the obstacle was removed and the agent either performed an inefficient but perceptually familiar action (left) or efficient but perceptually novel action (right) towards its goal.

plans to maximize rewards and minimize costs? Or do infants first analyze actions using leaner assumptions, such as the assumption that goal-directed actions will have certain perceptual features (e.g. that they move across flat supporting surfaces while facing their goals) and later acquire the principles guiding rational action?

Many experimental findings are consistent with the thesis that infants expect agents to behave rationally. In these experiments, infants first view an agent who moves on an efficient, curvilinear path to reach an object that stands behind an obstacle. Then the obstacle is removed, and infants are tested with the path that the agent had previously taken and a new, direct path. Infants' looking preferences provide evidence that they expect the novel, direct trajectory (Csibra, 2008; Csibra, Gergely, Bíró, Koós, & Brockbank, 1999; Gergely & Csibra, 2003; Gergely et al., 1995; Phillips & Wellman, 2005; Skerry, Carey, & Spelke, 2013) (Fig. 1). This expectation is early emerging (Skerry et al., 2013) and is applied broadly to both human-like (Phillips & Wellman, 2005; see also Gergely, Bekkering, & Király, 2002; Schwier, van Maanen, Carpenter, & Tomasello, 2006) and unfamiliar (Csibra, 2008; Gergely et al., 1995) agents. It is also inferentially powerful, licensing predictions about the configuration of an occluded physical scene (Csibra, Bíró, Koós, & Gergely, 2003) and the outcomes of ongoing actions (Csibra et al., 2003; Southgate & Csibra, 2009; Wagner & Carey, 2005).

Nevertheless, this family of findings is open to several interpretations. The richest construal is consistent with utility theory, based on representations of the relative costs of different actions and the relative rewards that these actions bring. Under this interpretation, infants, given two alternative actions with equal rewards and varying costs, expect agents to minimize the cost of their actions. However, at least three leaner interpretations are equally consistent with these findings. First, infants could construe agents as rational planners without a minimum function over costs: Infants could jointly rely on their assumptions about the solidity of agents and objects (Saxe et al., 2006), plus a set of general rules concerning the trajectory of motion agents follow when pursuing an unobstructed goal (e.g. that agents tend to move smoothly across supporting surfaces, face their goals, and move to them on straight paths), to generate this prediction. These assumptions about the features of actions could be innate, or learned, based on infants' past experiences performing their own actions and observing the actions of others. Alternatively, infants may have no initial expectations about the efficiency of agents, but may develop such expectations over the course of the experiment (e.g. by generalizing an agent's efficient behavior across changes in an its physical constraints). Lastly, infants could lack any ability to represent the cost of actions, but succeed in these experiments by generalizing perceptual features of an agent's actions from habituation to test (e.g. by learning that agents jump just high enough to clear the barrier[1]). Under the latter three interpretations, the content supporting infants' responses need not appeal to continuous, rich representations of cost.

1.2. Current experiments

Here, we test whether infants expect agents to minimize the cost of their actions against these alternatives. As reviewed above, the extant evidence for continuous representations of cost is consistent both with the broad and general principles of utility theory

---

[1] The last three interpretations could in principle explain findings from the control condition of past experiments (e.g. Csibra, 2008; Csibra et al., 1999; Gergely & Csibra, 2003; Gergely et al., 1995; Skerry et al., 2013), where the agent performs the same actions during habituation that are unconstrained by a barrier. This could (a) cause infants to suspend their predictions about an ostensibly irrational agent or (b) place a more difficult demand on them to learn the relation between the height of the jump and the height of the barrier (c.f. Csibra et al., 1999).

but also with narrower and more limited expectations, including the expectation that goal-directed agents travel to unobstructed goals by facing them and moving on straight paths, that agents who act rationally tend to continue to do so in the future, and that the path of an agent's action is predictable by the features of its environment.

To ask whether infants represent cost as a continuous variable, we present 3 experiments wherein infants' action predictions cannot be explained using rule-like assumptions about cost, learned from an agent's past efficiency, or learned from perceptual features of its past actions. We begin by testing infants for sensitivity to curved trajectories of motion that vary in efficiency. Whereas a system that represents the cost of actions as a set of local assumptions (e.g. agents move directly to their goals) would not distinguish between more or less efficient actions, a system that represents efficiency as a continuous variable would expect agents to minimize it. In Experiments 1 and 2, we test the hypothesis that 6-month-old infants expect agents to perform minimally costly actions when faced with a novel obstacle. In Experiment 3, we explore whether learning alone can account for this expectation by asking whether infants expect minimally costly action from an agent who previously engaged in inefficient actions.

2. Experiment 1

Experiment 1 tests whether infants discriminate between goal-directed actions over obstacles that vary in length and degree of curvature. When a barrier blocks an agent's direct path to a goal, do infants expect the agent to circumvent the barrier as efficiently as possible? If continuous representations of cost support expectations for efficient action, then infants should discriminate between the low and high trajectory of motion over the test barrier, and selectively recover attention when an agent takes a novel degree of cost given its new constraints.

2.1. Methods

2.1.1. Participants

Our sample included 20 full-term, healthy infants (10 female, $M_{age} = 5.95$ months, range = 5.57–6.63 months), comparable to studies of similar format and focus (e.g. Csibra, 2008; Csibra et al., 1999; Skerry et al., 2013). Seven more infants were tested, excluded and replaced (2 for fussiness that prevented study completion, 1 for failure to habituate, 1 for online coding error, 2 for technical failure, and 1 for parental interference). Sample size and exclusion criteria were fixed prior to the start of data collection, and decisions to exclude infants were made by researchers who were unaware of the order of events viewed by the infants. All participants were recruited from the greater Boston area and tested at the Laboratory for Developmental Studies at Harvard University with parental informed consent. Families received a small thank-you gift (e.g. a t-shirt or toy) for participating.

2.1.2. Materials

The animated events were created in Blender (Stichting Blender Foundation, 2016), synchronized with a custom audio track in iMovie, and presented using Keynote on an LCD projector screen 40″ in height and 52″ in width. Two speakers located on either side of the screen played all stimuli-related sounds. Infants' looking time data were coded online using Xhab64 (Pinto, 1995) software and offline using jHab (Casstevens, 2007).

2.1.3. Design

We used a habituation paradigm to probe infants' preferential looking towards two kinds of test events after reaching a predetermined habituation criterion (Fig. 2). All habituation and test trials began with an attention-getting star animation (2.0 s), and subsequently consisted of looped sequences of an animated event (5.0 s), which paused on the last frame (0.5 s) and was followed by a blank screen (0.5 s) before the next animation. Each animation featured a red spherical agent with eyes and a smiling mouth that began to move directly toward an unobstructed goal object (a blue cone), only to be impeded by a grey barrier that fell to rest with an audible thud between the agent and its goal. The agent then backed away, approached the obstacle, and leapt over it while making a popping sound, coming to rest next to the goal object. The timing of the jump was held constant across all habituation and test trials (0.9 s); thus, taller jumps were executed more rapidly.

In the *habituation events*, the height of the barrier varied across 3 levels (6, 5, and 3 Blender units) within each trial. The agent's jump height always aligned with that of the barrier (9.5, 8.5, and 6.5 units, respectively). Barriers and jumps of different heights were presented in a consistent pseudorandom order across trials.

The *test events* featured the same basic event structure with one critical change: a very short barrier (1 unit) obstructed the trajectory of the agent. To assess whether infants discriminate curved action trajectories on the basis of their efficiency with respect to this novel barrier, infants' attention was measured for two test trial types. On alternating test trials, the agent backed away from the barrier and performed either a low, efficient jump or a high, inefficient leap over it (4.5 and 9.5 units, respectively).

2.1.4. Procedure

Infants were seated on their caregivers' laps approximately 60″ away from the screen. Caregivers were instructed to keep their eyes closed and to refrain from interacting with their infants throughout the experiment, and were monitored for compliance.

After calibrating the infant to the screen using a squeaky toy, the researcher began the experiment. The researcher had access to a video feed of the infant's face, a computer screen indicating the current trial, and a third screen indicating when to conclude a trial and move from habituation to test. The researcher ran the experiment and coded looking time online while remaining unaware of the events displayed and test pair order, but could determine the precise start of each trial as well as the timing of the obstacle falling and the agent jumping over it based on auditory cues, which were identical in timing across all habituation and test trials.

Across both phases of the experiment, the experimenter began coding a trial immediately following the attention getter, and ended a trial once the infant had looked at the screen for 60 cumulative seconds or looked away for 2 consecutive seconds. The test phase began after infants' summed looking time from the most recent 3 habituation trials fell to below half their looking time in the first 3 trials (6–12 habituation trials total) and consisted of 3 pairs of test trials, order counterbalanced across subjects. These criteria were fixed prior to the start of data collection.

2.1.5. Coding and analyses

Videos of all test sessions were coded offline by observers without access to the events infants viewed, using the same thresholds as online coding, and reviewed for the predetermined subject exclusion criteria (fussiness that prevented study completion, online coding error, experimenter error, technical failure, parental interference, and failure to habituate). Further, if infants were determined to have missed a critical part of the test trial (i.e., never saw the agent jump over the test barrier), that test pair was marked and excluded from subsequent analyses. To assess the reliability of the offline-coded data, 100% of the test events were recoded independently by an additional researcher who was unaware of test pair order. The two coders agreed on the trial cutoffs

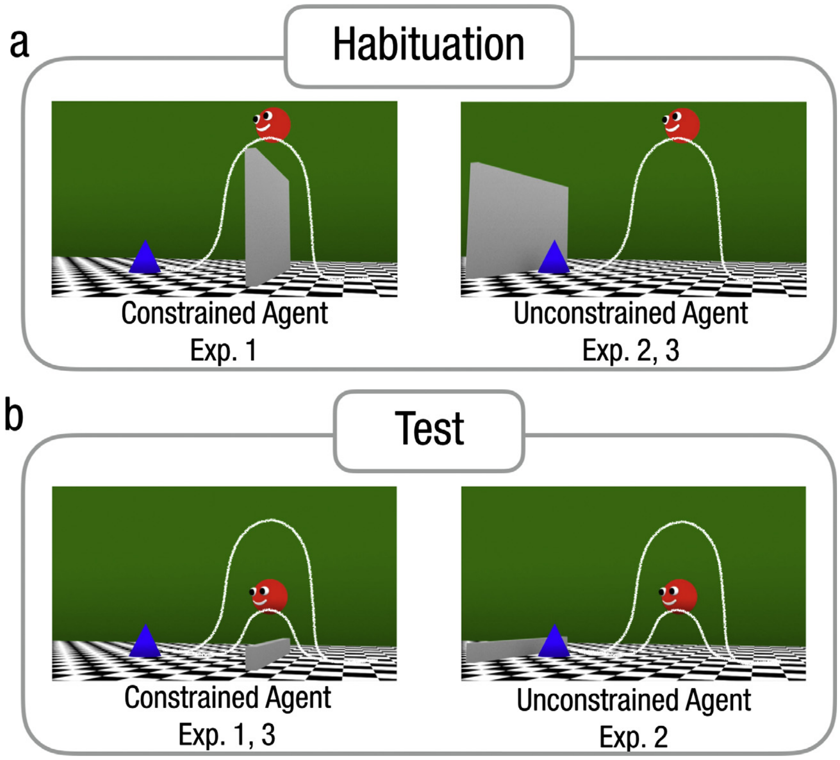

Fig. 2. Trial structure for Experiments 1–3, including (a) habituation to an agent leaping over tall barriers efficiently (left, Experiment 1) or performing identical motions without a physical constraint (right, Experiments 2 and 3) and (b) test, with the agent performing low and high jumps over a novel barrier (left, Experiments 1 and 3) or no barrier (right, Experiment 2). White lines indicate trajectories of motion.

for 94.17% of the test trials, and the intraclass correlation (ICC) between them was 0.969, 95% CI [0.957, 0.979]. Thus, the highly reliable primary offline coding data were used in our analyses.

Across all experiments, inferential statistics (e.g. model estimates, CIs) were fit to log-transformed looking time[2] (Csibra et al., 2016) averaged across all three test pairs, but plots and descriptive statistics feature raw looking times for ease of interpretation.

All analyses leveraged both traditional paired *t*-tests and linear mixed effects models in R (version 3.2.3; R Development Core Team, 2015). Linear mixed models were fit using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015). Detection of influential observations was conducted using the influence. ME package (Nieuwenhuis, te Grotenhuis, & Pelzer, 2012), a suite of methods for determining whether individual cases—in our models, participants—influenced the results such that their inclusion or exclusion could impact interpretation. Plots were produced using the ggplot2 package (Wickham, 2009). To explicitly take into account repeated measures, all mixed models included subject identity as a random intercept. Three classes of models were fit: (1) null models, featuring subject identity as the only predictor, (2) hypothesis-driven models, which included additional manipulated factor(s), and (3) exploratory models, which included additional non-hypothesis driven factors. We leveraged likelihood ratio tests (LRTs) to evaluate model fit by assessing whether the inclusion of certain predictors significantly reduced residual variance. All model-produced degrees of freedom were calculated using the Satterthwaite approximation method.

2.2. Results

2.2.1. Hypothesis-driven results

A hypothesis-driven model, including test action height (high versus low) as a fixed effect and subject identity as a random intercept, revealed that infants looked longer to the high test actions ($M = 16.24$ s, $SD = 12.54$) relative to the low test actions, ($M = 11.35$ s, $SD = 7.41$), 95% CI [0.106, 0.491], $B = 0.298$, $SE = 0.093$, $\beta = 0.462$, $t(20) = 3.191$, $p = 0.005$, two-tailed. A paired *t*-test supported this finding, 95% CI [0.098, 0.499], $t(19) = 3.110$, $d = 0.695$, achieved power = 0.838, $p = 0.006$, two-tailed. See Fig. 2. The hypothesis-driven model provided a better fit than a null model by a LRT, $\chi^2(1) = 8.229$, $p = 0.004$.

An analysis detecting influential cases using Cook's Distance (4/*n*, where *n* refers to the number of groups in the grouping factor in question; Van der Meer, te Grotenhuis, & Pelzer, 2010) revealed one influential subject (D = 0.201, cutoff = 0.2). Removal of this subject from the hypothesis-driven model produced an inferentially equivalent result, 95% CI [0.074, 0.448], $B = 0.261$, $SE = 0.097$, $\beta = 0.439$, $t(19) = 2.878$, $p = 0.010$, two-tailed.

2.2.2. Exploratory results

An exploratory model, testing for an fixed interactive effect of test presentation order and test action height, a fixed effect of sex, and subject identity as a random intercept, revealed no strong order effect, 95% CI [−0.185, 0.564], $B = 0.189$, $SE = 0.182$, $\beta = 0.239$, $t(20) = 1.039$, $p = 0.311$, two-tailed, or gender effects, 95% CI [−0.624, 0.429], $B = -0.097$, $SE = 0.256$, $\beta = -0.151$, $t(20) = -0.380$, $p = 0.708$, two-tailed. A LRT indicated that the exploratory model

---

[2] The log-normal distribution provided a better fit for raw LTs (log-likelihood = −376.95) across Exp. 1–3 than did the normal distribution (−416.10), maximum-likelihood fitting. We find inferentially equivalent results on hypothesis-driven tests by comparing the average proportion of time infants looked at the high test action against chance ($\mu = 0.5$), our original outcome measure, and using non-parametric analyses of raw looking times (see Supplemental Material available online). Our decision to present results using the dependent measure in the main text followed the recommendation of Csibra, Hernik, Mascaro, Tatone, and Lengyel (2016).

did not provide a better fit than the hypothesis-driven model, $X^2(3) = 1.863$, $p = 0.601$. No influential cases were detected.

2.3. Discussion

Experiment 1 provides evidence that is consistent with the hypothesis that infants leverage continuous representations rather than narrower assumptions about motion directness when reasoning about goal-directed action. Given a perceptually novel but efficient low jump and a perceptually familiar but inefficient high jump, infants recovered their attention to the inefficient jump, over and above the perceptual familiarity of this action. However, Experiment 1 alone cannot establish whether infants have a baseline looking preference for higher or faster motion, which could explain this finding in part or in whole. Experiment 2 explores this possibility.

3. Experiment 2

In Experiment 2, we followed the logic of past experiments (e.g. Gergely et al., 1995) to test for baseline looking preferences for higher or faster motion. We repeated Experiment 1 except for one critical change: All barriers fell *beyond* the goal object, such that the agent's actions were no longer physically constrained. If infants merely prefer faster or higher motion, then results should resemble those from Experiment 1, because the agent moved in identical ways across the two experiments. In contrast, if infants' responses in Experiment 1 are driven by representations of efficiency, their responses in Experiment 2 should differ, because all the actions in Experiment 2 were unconstrained and inefficient.

3.1. Methods

3.1.1. Participants

Our planned sample consisted of 20 full-term, healthy infants (10 female, $M_{\text{age}} = 6.10$ months, range = 5.70–6.67 months). An additional two infants were tested, excluded, and replaced due to online coding error.

3.1.2. Materials, design, and procedure

All aspects of the materials, design and procedure were identical to those from Experiment 1, except for the critical change in the location of the barrier from in between the agent and goal object to just beyond the goal object (Fig. 2).

3.1.3. Coding and analyses

The coding procedure was identical to that from Experiment 1. To test the reliability of the offline-coded data, 100% of the test events were recoded by an additional researcher who was unaware to the order of test trials. The two coders agreed on the trial cutoffs for 95.83% of the test trials, and the intraclass correlation (ICC) between them was 0.993, 95% CI [0.991, 0.995]. Thus, the highly reliable primary offline coding data were used in our analyses.

3.2. Results

3.2.1. Hypothesis-driven results

A hypothesis-driven model, including test action height (high versus low) as a fixed effect and subject identity as a random intercept, revealed that infants did not look longer to the high test event ($M = 11.21$ s, $SD = 6.04$) relative to the low test events ($M = 12.76$ s, $SD = 8.37$), 95% CI [−0.358, 0.235], $B = -0.071$, $SE = 0.149$, $\beta = -0.131$, $t(20) = -0.479$, $p = 0.637$, two-tailed. A paired t-test supported this finding, 95% CI [−0.390, 0.248], $t(19) = -0.466$, $d = 0.104$, achieved power = 0.073, $p = 0.646$, two-tailed. See Fig. 3. This model provided a fit no better than a null model by a LRT, $\chi^2(1) = 0.228$, $p = 0.633$.

An analysis detecting influential cases using Cook's Distance revealed one influential subject (D = 0.260, cutoff = 0.2). Removal of this subject from the hypothesis-driven model produced an inferentially equivalent result, 95% CI [−0.246, 0.280], $B = 0.017$, $SE = 0.128$, $\beta = 0.034$, $t(19) = 0.135$, $p = 0.894$, two-tailed.

3.2.2. Exploratory results

An exploratory model, testing for an fixed interactive effect of test presentation order and test action height, a fixed effect of sex, and subject identity as a random intercept revealed an order effect, 95% CI [0.232, 1.037], $B = 0.608$, $SE = 0.265$, $\beta = 1.122$, $t(20) = -2.297$, $p = 0.033$, two-tailed, and no gender effect, 95% CI [−0.354, 0.067], $B = -0.148$, $SE = 0.187$, $\beta = -0.273$, $t(20) = -0.789$, $p = 0.439$, two-tailed. A LRT indicated that the exploratory model did not provide a significantly better fit than the hypothesis-driven model, $\chi^2(3) = 5.315$, $p = 0.150$. Removal of one influential case (D = 0.382, cutoff = 0.2) produced an inferentially equivalent result, and this subject was removed from the following paired contrasts.

To probe the interactive effect between test order and test trial type, paired contrasts averaged across gender were extracted from the exploratory model and revealed that infants tended to look longer at whichever test event they first saw. Infants assigned to watch the high test event first looked longer at it ($M = 13.16$ s, $SD = 6.96$) than the low test event ($M = 10.39$ s, $SD = 5.16$), 95% CI [−0.120, 0.585], $t(21.24) = 1.371$, $p = 0.185$, two-tailed, and infants assigned to watch the low test event first looked longer at it ($M = 12.84$, $SD = 7.94$) than the high test event ($M = 9.606$, $SD = 4.63$), 95% CI [−0.594, 0.149], $t(21.24) = -1.242$, $p = 0.228$, two-tailed.

3.2.3. Comparing Experiments 1 and 2

To investigate the effect of barrier location (behind the goal object in Experiment 2, in front of it in Experiment 1) on the direction and extent to which infants discriminated between the test events, we fit a linear mixed effects model including an interactive fixed effect of action height (high versus low) and experiment (1 versus 2), plus a random intercept on subject identity. This analysis revealed that infants in Experiment 1 displayed a stronger looking preference than those in Experiment 2, 95% CI [0.017, 0.722,], $B = 0.369$, $SE = 0.176$, $\beta = 0.623$, $t(40) = 2.103$, $p = 0.042$, two-tailed. Removal of one influential case (D = 0.161, cutoff = 0.1) produced an inferentially equivalent result, 95% CI [−0.060, 1.026], $B = 0.281$, $SE = 0.157$, $\beta = 0.483$, $t(39) = 1.790$, $p = 0.081$, two-tailed.

3.3. Discussion

The findings of Experiment 2 are consistent with many previous reports that infants do not expect efficient action from an agent previously observed to move inefficiently (Csibra, 2008; Csibra et al., 1999; Gergely et al., 1995; Skerry et al., 2013; Southgate, Johnson, & Csibra, 2008). Infants who viewed events identical to those from Experiment 1, except for the position of the barrier, did not differentiate between the test events, indicating that infants in the previous experiment did not look longer at the inefficient action merely because it was higher or faster than the efficient one. So far, our findings are consistent with the hypothesis that infants expect agents to minimize the cost of their actions, rather than move smoothly and directly across supporting surfaces towards their goals.

Nevertheless, the interpretation of the above findings, including our own, is not clear. Two alternative construals remain: First, infants may learn over the course of the experiment that the agent will jump just high enough to clear the barrier and generalize this

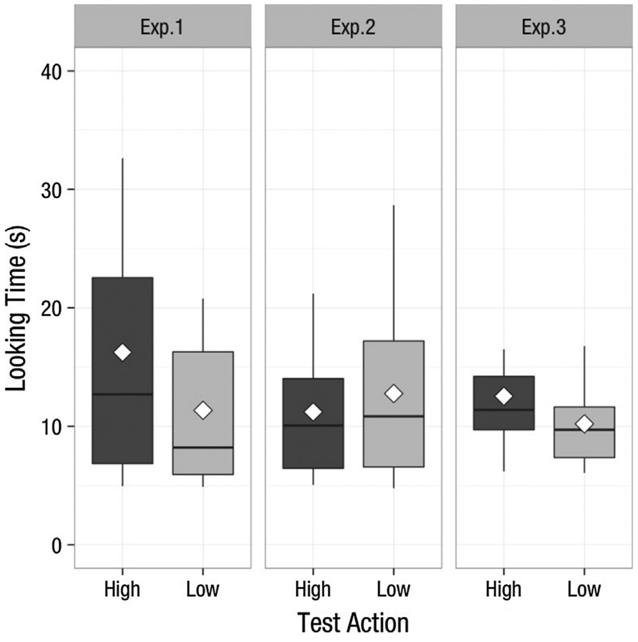

Fig. 3. Boxplots for raw looking times in seconds to test events in Experiments 1–3 ($N = 20$ per experiment). Boxes indicate interquartile ranges, bold horizontal lines indicate medians, and white diamonds indicate means.

relation at test, but fail to learn a similar relation when the barrier is away from the agent's path. Second, infants may not *a priori* expect agents to minimize the cost of their actions: they may expect previously efficient agents to continue acting efficiently, but when shown an agent behaving inefficiently in one context, suspend all predictions about its subsequent actions (Csibra, 2003; Csibra et al., 1999; Gergely et al., 1995; Southgate et al., 2008). In summary, a key remaining question is whether infants have an overhypothesis (Goodman, 1983) that agents minimize the cost of their actions. Experiment 3 was undertaken to address this question.

4. Experiment 3

Experiment 3 investigates whether infants expect minimally costly action given no prior evidence of an agent's rational action or opportunity to learn about the trajectory of the action given the barrier as a point of reference. If infants assume an overhypothesis on observing minimally costly actions, then observing inefficient action may suspend the expectations of efficiency only in the narrow context in which the agent is acting: infants may continue to expect the agent to act efficiently in new situations. To test for this possibility, we paired the habituation events from Experiment 2, in which the agent acts efficiently, with the test events of Experiment 1, in which the agent confronts a new obstacle. If infants expect agents to act efficiently only when they have prior evidence of its efficiency, or if infants are merely adept at learning that an agent will jump just as high as necessary to clear the barrier, then they should hold no expectations here. In contrast, if infants represent the principle of efficiency as an overhypothesis, then they may expect a previously non-goal-directed agent to minimize the cost of its action the very first time it faces a physical obstacle.

4.1. Methods

4.1.1. Participants

Our planned sample included 20 full-term, healthy infants (10 female, $M_{age} = 5.84$ months, range = 5.40–6.13 months). An additional 5 infants were tested, excluded and replaced (2 for fussiness and 2 for online coding error, and 1 for missing a critical portion of the events for all 3 test pairs).

4.1.2. Materials, design, and procedure

All aspects of the materials, design and procedure were identical to those from Experiment 1 and 2 except for the configuration of the habituation and test events. To test whether infants expect an agent to navigate over a low constraint efficiently without ever having seen the agent act in a goal-directed manner, we paired the habituation events from Experiment 2 with the test events from Experiment 1. That is, infants were habituated to an agent that performed unconstrained actions, and then were tested on events in which a constraint was in the agent's way for the very first time. See Fig. 2.

4.1.3. Coding and analyses

The coding procedure was identical to that from Experiments 1 and 2. To assess the reliability of the offline-coded data, 100% of the test events were recoded by an additional researcher who was unaware of test pair order. The two coders agreed on the trial cutoffs for 94.17% of the test trials, and the intraclass correlation (ICC) between them was 0.972, 95% CI [0.960, 0.980]. Thus, the highly reliable primary offline coding data were used in our analyses.

4.2. Results

4.2.1. Hypothesis-driven results

A hypothesis-driven model, including test action height (high versus low) as a fixed effect and subject identity as a random intercept, revealed that infants looked longer to the high test event ($M = 12.54$ s, $SD = 5.01$) relative to the low test event ($M = 10.20$ s, $SD = 3.56$), 95% CI [0.056, 0.325], $B = 0.190$, $SE = 0.066$, $\beta = 0.522$, $t(20) = 2.906$, $p = 0.009$, two-tailed. A paired t-test supported this finding, 95% CI [0.050, 0.331], $t(19) = 2.833$, $d = 0.633$, achieved power = 0.766, $p = 0.011$, two-tailed. This model outperformed a null model by a LRT, $X^2(1) = 7.046$, $p = 0.008$.

An analysis detecting influential cases using Cook's Distance revealed one influential subject (D = 0.256, cutoff = 0.2). Removal of this subject from the hypothesis-driven model produced an inferentially equivalent result, 95% CI [0.034, 0.275], $B = 0.155$, $SE = 0.058$, $\beta = 0.452$, $t(19) = 2.649$, $p = 0.016$, two-tailed.

4.2.2. Exploratory results

An exploratory model testing for order and sex effects included fixed interactive effect of test presentation order and test action height, a fixed effect of sex, and subject identity as a random intercept. This analysis revealed neither an order effect, 95% CI [−0.207, 0.239], $B = 0.061$, $SE = 0.130$, $\beta = 0.167$, $t(20) = 0.468$, $p = 0.645$, two-tailed, nor an effect of sex, 95% CI [−0.169, 0.401], $B = 0.116$, $SE = 0.138$, $\beta = 0.318$, $t(20) = 0.838$, $p = 0.412$, two-tailed. A LRT indicated that the exploratory model did not provide a significantly better fit than the hypothesis-driven model, $\chi^2(3) = 0.910$, $p = 0.823$. Removal of one influential case (D = 0.218, cutoff = 0.2) produced an inferentially equivalent result.

4.2.3. Comparing Experiments 1 and 3

To investigate the effect of the initial behavior of the agent during habituation (efficient in Experiment 1 and inefficient in Experiment 3) on infants' responses to subsequent constrained actions during test, we fit a model including interactive fixed effect between study (Experiment 1 versus 3) and test action height (high versus low) and a random intercept for subject identity. This analysis revealed no consistent difference in looking preference across the 2 experiments, 95% CI [−0.337, 0.121], $B = -0.108$, $SE = 0.114$, $\beta = -0.207$, $t(40) = -0.945$, $p = 0.350$, two-tailed. Removal of two influential cases (D = 0.143 and 0.101, cutoff = 0.1) produced an inferentially equivalent result, 95% CI [−0.328, 0.095], $B = -0.117$, $SE = 0.105$, $\beta = -0.240$, $t(38) = -1.110$, $p = 0.274$, two-tailed. An additional model removing this interaction revealed that infants looked longer to the high test action collapsing across both experiments, 95% CI [0.128, 0.360], $B = 0.244$, $SE = 0.058$, $\beta = 0.468$, $t(40) = 4.234$, $p < 0.001$, two-tailed, but not differently across Experiments 1 and 3 collapsing across trial type, 95% CI [-0.340, 0.256], $B = -0.042$, $SE = 0.148$, $\beta = -0.080$, $t(40) = -0.283$, $p = 0.779$, two-tailed. Removal of influential cases in the second model (D = 0.158 and 0.108, cutoff = 0.1) produced inferentially equivalent results: Infants selectively recovered attention to an inefficient action performed over a novel barrier, 95% CI [0.116, 0.0.345], $B = 0.231$, $SE = 0.057$, $\beta = 0.504$, $t(38) = 4.045$, $p < 0.001$, two-tailed, and did so regardless of whether they previously saw it act efficiently, 95% CI [−0.181, 0.342], $B = 0.081$, $SE = 0.130$, $\beta = 0.176$, $t(38) = 0.618$, $p = 0.540$, two-tailed. See Supplemental Material available online for additional analyses across experiments.

4.3. Discussion

After an action-relevant change in the location of an obstacle, infants expected a previously inefficient agent to minimize cost in Experiment 3, as in Experiment 1. This result contrasts with the findings of previous studies (Csibra, 2008; Csibra et al., 1999; Gergely et al., 1995; Phillips & Wellman, 2005; Skerry et al., 2013) as well as Experiment 2, in which infants saw no change in the physical constraints of the agent between habituation and test. This finding shows that by 6 months of age, infants expect agents to minimize the cost of their actions under conditions where learning about the efficiency of or perceptual regularities found its actions was not possible. Thus, infants appear to assume an overhypothesis on observing minimally costly action given a change in an agent's constraints (Experiment 1) or the first time a constraint is introduced (Experiment 3).

5. General discussion

Three experiments provided evidence that by 6 months of age, infants represent the principle of efficiency as an expectation that agents minimize the cost of their actions. When presented with action trajectories differing in curvature, infants differentiated between these actions on the basis of their efficiency, over and above perceptual differences in height or velocity. This finding indicates that the principle of efficiency is not articulated only by local assumptions about agents and their actions, such as the assumption that agents move in straight lines or directly towards their goals. Its activation also does not depend on prior observation of efficient action: infants even applied this expectation to an agent whose previous actions were all inefficient, indicating that their responses cannot be explained by learning, during the experiment, either that an agent acts efficiently or that the height of its jumps bears a consistent geometric relationship to the height of the barrier that it jumps over. We suggest that this expectation is carried in an overhypothesis (Goodman, 1983) on minimal costs. That is, infants have an inductive bias to expect maximally efficient action from agents situated in new physical contexts, even if they never acted efficiently in past contexts. This assumption may guide their analysis and learning about the social world by biasing their expectations towards observing rational behavior.

Though our experiments only address representations of cost, their results are consistent with the thesis that early action understanding is expressed as a richer system of reasoning about how hidden variables like effort, desire, and belief guide action planning. Nevertheless, it is still an empirical question how richly utility theory articulates early understanding of action. We conclude by describing three future lines of research that bear on this question.

First, although we argue that infants hold an overhypothesis on minimal costs, it is not clear how rich and abstract the variables are that enter into these computations. It is possible, for example, that infants' understanding of cost is restricted to one or a few dimensions of action: In Experiments 1 and 3, infants may have leveraged a minimum function on the length or indirectness of the agent's actions without considering the psychological cost of planning and the physical cost executing them. Further research could reveal whether infants' intuitions about cost are best described in terms of these leaner assumptions, or instead in terms of the physical work required to execute actions and the mental effort required to plan them. If infants have a general, abstract assumption that agents minimize cost, then they might expect agents to choose an action that requires less force or less planning under conditions where features like path length are held equal.

Second, the present research raises questions concerning the inferential role of costs within a broader schema of action understanding. Do infants, like older children, expect agents to plan utility-maximizing actions, considering not only the costs of different actions but also the rewards that these actions bring (Jara-Ettinger et al., 2015)? To our knowledge, no research reveals whether infants reason about costs and rewards by representing a function that subsumes both roles. Future experiments could test, for example, whether infants can infer an agent's desires and beliefs about the world given the degree of effort it expends.

Finally, these results do not reveal how this knowledge is acquired and how it develops over the first six months. How does a cognitive system come to represent functions over abstract costs and identify the range of events to which such functions apply? According to one theoretical stance, concepts like goals are constructed from sensorimotor mappings between observed and experienced actions (Paulus, 2012; Paulus, Hunnius, Vissers, & Bekkering, 2011; Woodward, 2009). According to a second account, these concepts are embedded in an innate, fully productive schema for action understanding that supports representations of the actions we experience and observe in the world (Carey, 2009; Gergely & Csibra, 2003). Recent evidence suggests that sensitivity to the costs of actions does not rely on action experience alone (Skerry et al., 2013), but the precise role of experience in action understanding is not known. A third possibility is that infants begin with a skeletal set of assumptions about agents and their actions, which is then enriched in in the first years of life, in line with their developing knowledge about the physical world (Baillargeon, 2002). Under this account, infants begin with some assumptions about costs and rewards, but face the challenges of learning the specific costs and rewards of actions and states in the world and constructing the form of knowledge that best captures how agents plan behavior. Research documenting the ontogeny and phylogeny of action understanding could distinguish these possibilities.

5.1. Conclusion

Our intuitive psychology is supported by the assumption that agents plan actions so as to maximize desired states of the world (rewards) while minimizing effort (costs). Characterizing the functions, variables, and procedures that articulate these assumptions not only constrains our theories of mature social cognition but also provides a framework under which we investigate its changes over development. Here, we applied this approach to probe the representations that support expectations for efficient action in the first year of life. We discovered that 6-month-old infants use a minimum function over costs to guide their expectation for rational action, and that they apply this expectation even to agents whose previous actions were inefficient. Our case study provides a step toward characterizing the early cognitive substrates on which humans build a rich, abstract, and productive system for action understanding.

Acknowledgements

The authors thank the families who volunteered to participate, A. Aguirre, C. Kerwin, and T. Ladd for research assistance, R. Guzman and N. Kalra for administrative assistance, and to S. Carey, J. Tenenbaum, T. Ullman, and two anonymous reviewers for comments and advice. This material is based upon work supported by the Center for Brains, Minds and Machines (CBMM) funded by National Science Foundation STC award CCF-1231216, and by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1144152.

Appendix A. Supplementary material

All data and materials have been made publicly available via the Open Science Framework, and can be accessed at http://osf.io/sxdtg/ and http://osf.io/4qw45/, respectively. Supplementary data associated with this article can also be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2016.12.007.

References

Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: A summary in eight lessons. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (1st ed., pp. 47–83). Blackwell Publishers Ltd.. http://dx.doi.org/10.1002/9780470996652.ch3.

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition, 113*(3), 329–349. http://dx.doi.org/10.1016/j.cognition.2009.07.005.

Bartsch, K., & Wellman, H. (1989). Young children's attribution of action to beliefs and desires. *Child Development, 60*(4), 946–964.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1). http://dx.doi.org/10.18637/jss.v067.i01.

Carey, S. (2009). *The origin of concepts*. New York, NY: Oxford University Press.

Casstevens, R. M. (2007). jHab: Java Habituation Software (Version 1.0.0) [Computer software]. Chevy Chase, MD.

Csibra, G. (2003). Teleological and referential understanding of action in infancy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 358*(1431), 447–458. http://dx.doi.org/10.1098/rstb.2002.1235.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition, 107*(2), 705–717. http://dx.doi.org/10.1016/j.cognition.2007.08.001.

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science, 27*(1), 111–133. http://dx.doi.org/10.1016/S0364-0213(02)00112-X.

Csibra, G., Gergely, G., Bíró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of "pure reason" in infancy. *Cognition, 72*(3), 237–267. http://dx.doi.org/10.1016/S0010-0277(99)00039-6.

Csibra, G., Hernik, M., Mascaro, O., Tatone, D., & Lengyel, M. (2016). Statistical treatment of looking-time data. *Developmental Psychology, 52*(4), 521–536. http://dx.doi.org/10.1037/dev0000083.

Csibra, G., & Volein, A. (2008). Infants can infer the presence of hidden objects from referential gaze information. *British Journal of Developmental Psychology, 26*(1), 1–11. http://dx.doi.org/10.1348/026151007X185987.

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

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature, 415*(6873), 755. http://dx.doi.org/10.1038/415755a.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences, 7*(7), 287–292. http://dx.doi.org/10.1016/S1364-6613(03)00128-1.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*(2), 165–193. http://dx.doi.org/10.1016/0010-0277(95)00661-H.

Goodman, N. (1983). *Fact, fiction, and forecast*. New York, NY: Bobbs-Merrill (Original work published 1955).

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences, 20*(8), 589–604. http://dx.doi.org/10.1016/j.tics.2016.05.011.

Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition, 140*, 14–23. http://dx.doi.org/10.1016/j.cognition.2015.03.006.

Luo, Y., & Johnson, S. C. (2009). Recognizing the role of perception in action at 6 months. *Developmental Science, 12*(1), 142–149. http://dx.doi.org/10.1111/j.1467-7687.2008.00741.x.

Muentener, P., & Carey, S. (2010). Infants' causal representations of state change events. *Cognitive Psychology, 61*(2), 63–86. http://dx.doi.org/10.1016/j.cogpsych.2010.02.001.

Nieuwenhuis, R., te Grotenhuis, M., & Pelzer, B. (2012). Influence. ME: Tools for detecting influential data in mixed effects models. *R Journal, 4*(2), 38–47.

Paulus, M. (2012). Action mirroring and action understanding: an ideomotor and attentional account. *Psychological Research, 76*(6), 760–767. http://dx.doi.org/10.1007/s00426-011-0385-9.

Paulus, M., Hunnius, S., Vissers, M., & Bekkering, H. (2011). Imitation in infancy: Rational or motor resonance? *Child Development, 82*(4), 1047–1057. http://dx.doi.org/10.1111/j.1467-8624.2011.01610.x.

Phillips, A. T., & Wellman, H. M. (2005). Infants' understanding of object-directed action. *Cognition, 98*(2), 137–155. http://dx.doi.org/10.1016/j.cognition.2004.11.005.

Pinto, J. (1995). *Xhab64 [computer software]*. Palo Alto, CA: Stanford University.

R Development Core Team (2015). R: A language and environment for statistical computing (Version 3.2.1) [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <http://www.R-project.org/>.

Saxe, R., Tenenbaum, J. B., & Carey, S. (2005). Secret agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science, 16*(12), 995–1001. http://dx.doi.org/10.1111/j.1467-9280.2005.01649.x.

Saxe, R., Tzelnic, T., & Carey, S. (2006). Five-month-old infants know humans are solid, like inanimate objects. *Cognition, 101*(1), B1–B8. http://dx.doi.org/10.1016/j.cognition.2005.10.005.

Schwier, C., van Maanen, C., Carpenter, M., & Tomasello, M. (2006). Rational imitation in 12-month-old infants. *Infancy, 10*(3), 303–311. http://dx.doi.org/10.1207/s15327078in1003_6.

Skerry, A. E., Carey, S. E., & Spelke, E. S. (2013). First-person action experience reveals sensitivity to action efficiency in prereaching infants. *Proceedings of the National Academy of Sciences of the United States of America, 110*(46), 18728–18733. http://dx.doi.org/10.1073/pnas.1312322110.

Southgate, V., & Csibra, G. (2009). Inferring the outcome of an ongoing novel action at 13 months. *Developmental Psychology, 45*(6), 1794–1798. http://dx.doi.org/10.1037/a0017197.

Southgate, V., Johnson, M. H., & Csibra, G. (2008). Infants attribute goals even to biomechanically impossible actions. *Cognition, 107*(3), 1059–1069. http://dx.doi.org/10.1016/j.cognition.2007.10.002.

Stichting Blender Foundation (2016). Blender (Version 2.78) [Computer software]. Retrieved from <https://www.blender.org/download/>.

Van der Meer, T., te Grotenhuis, M., & Pelzer, B. (2010). Influential cases in multilevel modeling: A methodological comment. *American Sociological Review, 75*, 173–178.

Wagner, L., & Carey, S. (2005). 12-Month-old infants represent probable endings of motion events. *Infancy, 7*(1), 73–83. http://dx.doi.org/10.1207/s15327078in0701_6.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655–684.

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*(1), 1–34. http://dx.doi.org/10.1016/S0010-0277(98)00058-4.

Woodward, A. L. (2009). Infants' grasp of others' intentions. *Current Directions in Psychological Science, 18*(1), 53–57. http://dx.doi.org/10.1111/j.1467-8721.2009.01605.x.