Toddler's action prediction: Statistical learning of continuous action sequences

Claire D. Monroy [a,b], Sarah A. Gerson [a,c], Sabine Hunnius [a]

[a] Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen
[b] Ohio State University Wexner Medical Center
[c] Cardiff University, School of Psychology

In Press, Journal of Experimental Child Psychology, 2017

Abstract

The current eye-tracking study investigated whether toddlers use statistical information to make anticipatory eye movements while observing continuous action sequences. In two conditions, 19-month-old participants watched either a person performing an action sequence (Agent condition) or a self-propelled visual event sequence (Ghost condition). Both sequences featured a statistical structure in which certain action pairs occurred with deterministic transitional probabilities. Toddlers learned the transitional probabilities between the action steps of the deterministic action pairs and made predictive fixations to the location of the next action in the Agent condition, but not the Ghost condition. These findings suggest that young toddlers gain unique information from the statistical structure contained within action sequences, and are able to successfully predict upcoming action steps based on this acquired knowledge. Further, predictive gaze behavior was correlated with reproduction of sequential actions following exposure to statistical regularities. This study extends prior developmental work by showing that statistical learning can guide the emergence of anticipatory eye movements during observation of continuous action sequences.

Keywords: action prediction, action observation, statistical learning, cognitive development, eye-tracking

Toddlers' Action Prediction: Statistical Learning of Continuous Action Sequences

1.0 Introduction

Action sequences are continuous, complex streams of movement. Underlying this intricate flow of information, actions contain statistical regularities that provide information about the structure of observed behavior. For instance, a typical experience for an infant might be observing a parent bake a pie by first gathering ingredients, preparing dough, mixing the filling, and placing it in the oven. This action sequence includes many movements with unique kinematics, a temporal structure, and contextual information. A complex challenge for infants is to efficiently process this diverse and continuous flow of motion, recognize and predict the sub-steps that define the overarching action, and reach an understanding of the overall goal of the behavior.

Recent developmental studies have demonstrated that action prediction—the ability to perform predictive gaze shifts during action observation—is mediated by infants' motor capabilities and their ability to associate actions with their effects (Falck-Ytter, Gredebäck, & von Hofsten, 2006; Hunnius & Bekkering, 2014; Jovanovic, Kiraly, Elsner, Gergely, Prinz, & Aschersleben, 2007; Paulus, Hunnius, van Wijngaarden, Vrins, van Rooij, & Bekkering, 2011; Reid, Hoehl, Grigutsch, Groendahl, Parise, & Striano, 2009; Stapel, Hunnius, van Elk, & Bekkering, 2010; Woodward & Sommerville, 2000). One outstanding question is whether statistical learning skills also support action prediction early in development. Statistical learning, defined as sensitivity to statistical regularities in the environment, has been shown to be a powerful learning tool that emerges early in development (Ruffman, Taumoepeau, & Perkins, 2012; Saffran, Aslin, & Newport, 1996). Infants' and toddlers' statistical learning skills were initially investigated in the auditory domain, and recently researchers have shown that statistical learning also allows infants to extract structure from visual input as well (Kirkham, Slemmer & Johnson, 2002). The present study aimed to determine whether toddlers can discover the statistical regularities in action sequences and accurately predict upcoming actions and their effects during online observation.

1.1 Statistical learning: sensitivity to sequential structure

The transitional probabilities (TPs) that define the order of events in a sequence determine how predictable an upcoming event is relative to other possible events. Previous research has demonstrated that young infants and children are sensitive to probabilistic information in different sensory domains, including in the action domain (e.g., Amso & Davidow, 2012; Buchsbaum, Gopnik, Griffiths, & Shafto, 2011; Canfield & Haith, 1991; Wentworth, Haith, & Hood, 2002). These statistical learning paradigms typically feature a learning phase containing sequences with a consistent structure such that certain elements always occur in a predictable sequential order. Infants and toddlers subsequently demonstrate a novelty response by looking longer to test sequences that deviate from the structure observed in the learning phase (e.g., Baldwin, Baird, Saylor, & Clark, 2001; Saylor, Baldwin, Baird, & LaBounty, 2007).

For example, in a study by Baldwin and colleagues (2001), 10-month-olds were familiarized to videos of daily actions such as sweeping. In a subsequent test phase, they watched versions of these videos in which pauses were inserted in the midst of the actor’s motion (‘interrupting’ videos), or at natural breakpoints in the action stream (‘intact’ videos). Infants increased visual attention to the interrupting videos, which was interpreted as evidence that they perceived a violation in the expected structure. Additional research has shown that, rather than being restricted to highly familiar events, infants can also learn the structure of novel, abstract sequences (e.g., the dancing of a starfish) in which the only cues for segmentation are the TPs between dynamic events (Stahl, Romberg, Roseberry, Golinkoff, & Hirsh-Pasek, 2014).

Research on the role of statistical structure in action processing has until recently focused on segmentation abilities via reactive measures of learning. The current experiment investigated whether toddlers can anticipate upcoming actions or events during online learning of sequential actions using their statistical learning skills. Recently, Romberg and Saffran (2012) and Tummeltshammer and Kirkham (2013) demonstrated that infants can learn to make anticipatory fixations to the location where a visual shape will appear following an auditory or visual cue. Although this work suggests that infants can learn visual spatiotemporal relations and predict upcoming locations, it is unknown whether they can anticipate the occurrence of a particular action or visual event, independent of its location. In the current study, toddlers could not simply learn object-location or audiovisual associations, but needed to learn the predictable transitions between successive actions or events to correctly anticipate what would occur next. Though visual SL abilities are generally considered to be domain-general (e.g., Kirkham et al., 2002), prior research has also shown that SL is subject to some constraints and facilitative effects depending on the nature and complexity of the observed input (e.g., Krogh, Vlach, & Johnson, 2013). In the current study, we hypothesized that toddlers would more readily learn statistical structure contained within observed action sequences relative to (non-action) ghost sequences. In addition to general perceptual and attentional processes, perceiving human action engages unique cognitive mechanisms (Schubotz, 2007) that may extend to how statistical regularities in action sequences are learned, as we discuss in the following section.

1.2 Relations between action prediction and execution in development

A growing body of evidence suggests that action prediction is constrained by infants' developing action repertoire. The ability to predict action outcomes relates to developing motor capabilities or experiences (e.g., Gredebäck & Kochukhova, 2009; Kanakogi & Itakura, 2011; Kochukhova & Gredebäck, 2010; Stapel, Hunnius, Meyer, & Bekkering, 2016). These developmental findings are consistent with embodied accounts of action observation, which propose that action prediction arises from activated representations in the observer's own motor system (Adams, Shipp, & Friston, 2013; Elsner, D'Ausilio, Gredebäck, Falck-Ytter, & Fadiga, 2013; Flanagan & Johansson, 2003; Hunnius & Bekkering, 2014; Keysers, Kaas, & Gazzola, 2010). According to these embodied accounts, observing actions engages the observer's own motor system and provides an additive learning benefit beyond other forms of observational learning.

The 'ghost display' method is a technique for isolating the unique role of action, relative to other visual information (coined by Fawcett, Skinner, & Goldsmith, 2002). In ghost displays, objects move on their own without an agent performing motor acts, thus preserving critical components of the events such as motion and object cues without the action cue provided by a hand or person (for a review, see Hopper, 2010). Young children are more successful at learning and imitating when observing actions compared to ghost displays (Hopper, Lambeth, Schapiro, & Whiten, 2014; Thompson & Russell, 2004). In the current experiment, we adopted the ghost-display method to examine whether action observation uniquely facilitates prediction of sequential actions.

1.3 Actions and their effects

Actions result in sensory consequences, such as the sound of a whistle when we blow it. According to the *action-effect principle* (Prinz, 1997), action effects are central to action prediction and planning (Elsner et al., 2002). Empirical evidence has shown that effects influence infants' processing of observed actions and their goals (Jovanovic et al., 2007). Twelve-month-olds, for instance, will imitate an action followed by an effect more frequently than other observed actions (Klein, Hauf, & Aschersleben, 2006), particularly when they observe an agent cause the effect relative to a non-social animation that precedes the same effect (Danish & Russell, 2007). This suggests that toddlers may be more likely to make predictions via exposure to statistical regularities particularly in the case when an *agent* produces an action with a sensory *effect*. We test this hypothesis by manipulating whether or not deterministic pairs result in an action-effect within each condition (Agent and Ghost).

In a recent study, two-year-olds observed probabilistic relations between actions and effects and subsequently performed only the actions that were more likely to produce the effect (Waismeyer, Meltzoff, & Gopnik, 2015). The authors explicitly demonstrated the actions in isolated sequences in a live setting and encouraged toddlers to try to reproduce the effect. In everyday situations, learning opportunities are not always so explicit and sometimes action outcomes are not necessarily very salient. Would young children similarly reproduce action contingencies implicitly embedded within continuous sequences? We investigated toddler's action reproduction following implicit learning as an additional exploratory measure in the current experiment.

1.4 The current study

Our central question was whether toddlers spontaneously exploit statistical information in continuous action sequences to predict upcoming actions. We focused on three central components of action processing: (a) the role of observing actions vs. visual events (Agent and Ghost conditions; between-subjects), (b) the influence of action effects vs. lack of effects (Effect and No-effect pairs; within-subjects), and (c) the correlation between predictive behaviors during observation and subsequent action production. For (a) and (b), we measured predictive gaze during learning using an anticipatory fixation paradigm, which has been established as a measure of visual predictions (Hunnius & Bekkering, 2010; Romberg & Saffran, 2012; Tummeltshammer & Kirkham, 2013). For (c), we analyzed the action sequences that toddlers produced post-observation and correlated these with their predictive looking during learning.

2.0 Methods

2.1 Participants

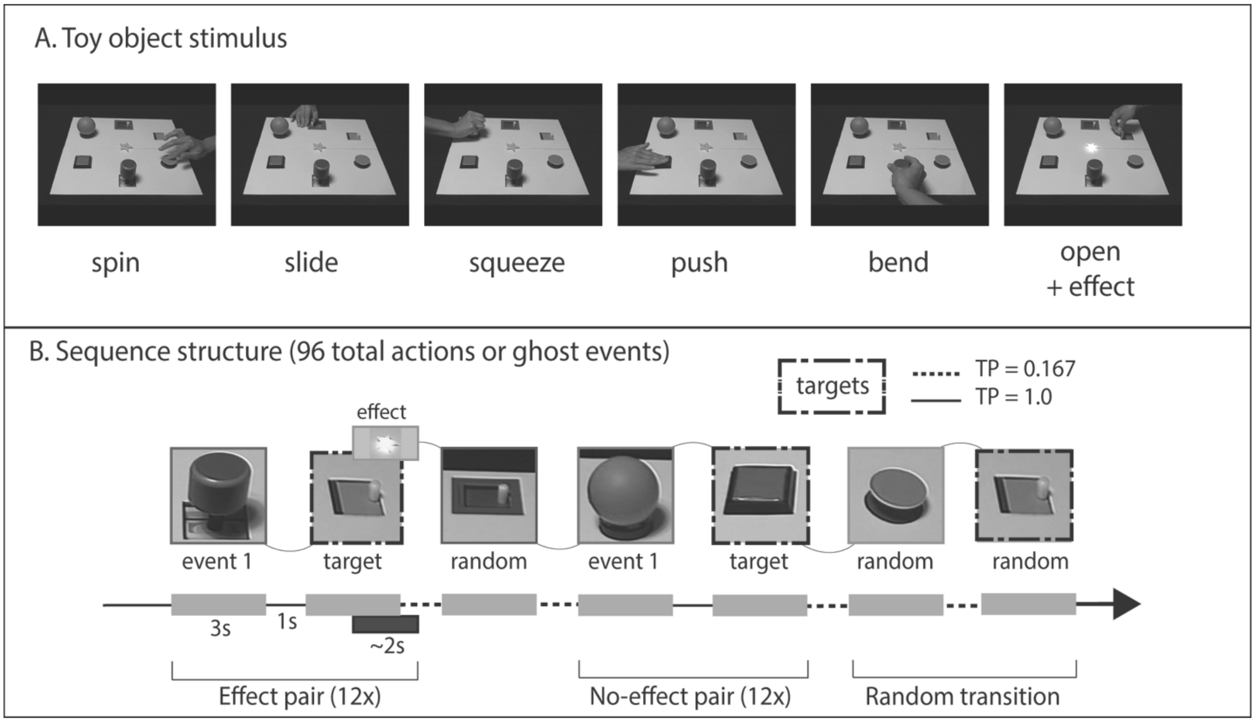
Eighty-six toddlers participated in this study. Twenty-seven toddlers were excluded due to insufficient gaze data (see Analysis section), excessive fussiness or failure of the eye-tracker. Fifty-nine toddlers remained in the final sample (mean age = 19.1 months, range: 18.5-20.5 months, *SD* = .82, 23 females): 31 in the Agent and 28 in the Ghost condition. Written consent was acquired from all parents, and families received a small gift or 10 Euros for participation. Participants were recruited from a database of families that reflects the demographics of the surrounding region (primarily Caucasian and middle-class).

2.2 Stimuli

Toddlers observed a full-screen (1280x1024 pixels) film of a sequence featuring a multi-object toy. This toy contained six unique objects, with distinct manual affordances, and a central star-shaped light (Figure 1A). Movies were filmed with a Sony HandyCam video camera and edited using Adobe Premiere Pro Cs5 software. The same toy was presented to toddlers during play sessions.

Eye movements were recorded using a Tobii T60 eye-tracker (Tobii, Stockholm, Sweden) with a 17” monitor (distance from screen: 63cm; visual angle: 148.9°). Stimuli were presented with Tobii ClearView AVI presentation software and sounds were played through external speakers.

Figure 1. A: Selected frames illustrating each action in the Agent condition; B: Schematic illustrating the sequence structure. Gray rectangles below the objects depict the timeline; the action-effect overlapped in time with the second action of the Effect pair. Pairs and random events could be followed by any of the six possible events. TP = transitional probability.



2.2.1 Sequence

We created four pseudo-randomized sequences of 96 individual events using the program Mix (van Casteren & Davis, 2006). These sequences contained two deterministic pairs (TP = 1.0). The remaining sequence events occurred randomly (TP = .167 or 1/6; Figure 1B). For clarity, we will refer to the second events of the deterministic pairs as *targets*, as these were the events that became predictable as the sequence unfolded. One deterministic pair caused a central star on the toy to light up (Effect pair), and the second pair had no effect (No-effect pair). The effect occurred at a natural mid-point of the target action or event. For example, during *open*, the light turned on the moment the yellow door was fully open and turned off after it closed. The objects corresponding to Effect and No-effect pairs were counterbalanced across toddlers.

Importantly, both targets could also occur at random elsewhere in the sequence (see Figure 1B). In these instances, the action-effect did not occur. This was to ensure that the Effect pair target was not independently associated with the effect. It also allowed us to measure whether toddlers made visual predictions to the same target in different probabilistic contexts (deterministic vs random, see section *3.3*). Additional constraints were imposed such that no single event could occur more than three times consecutively. All events occurred 12 times (except for targets, which appeared 12 times within pairs and 12 times at random). Thus, participants viewed 24 deterministic pairs (12 each of Effect and No-effect pairs) and 48 random unpaired events in the total sequence of 96 individual events.

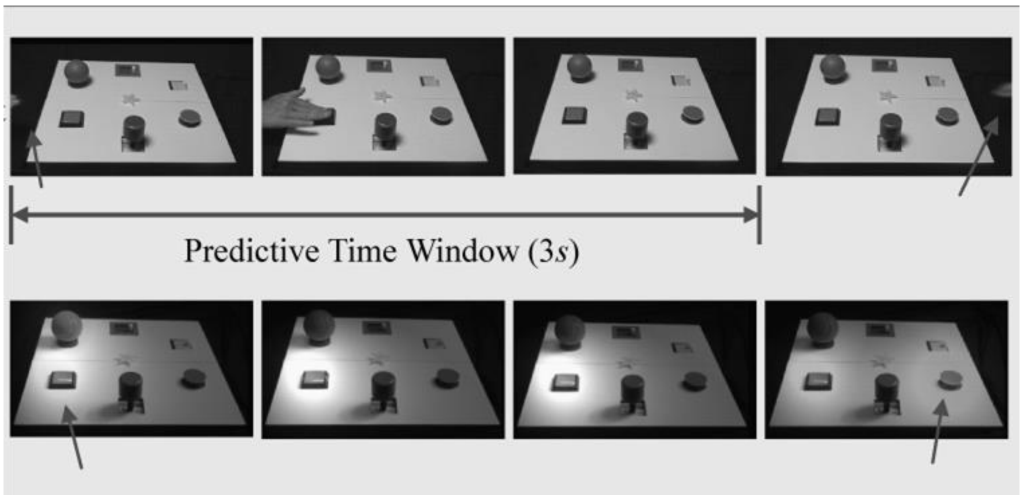
Sequences were divided into four blocks, with the viewing orientation rotated during each block to ensure that the toddler could not predict the next action or event from its location on the screen. Each block lasted approximately 90s and consisted of 24 events. Brief animations were presented in the center of the screen between blocks to engage and maintain attention. At the beginning of a block, a still frame was presented (4s) to allow toddlers to reorient to the new perspective. Engaging music was played throughout the films, which did not correspond in any way to the unfolding action sequence.

2.2.2 Agent Condition

In the Agent condition, an actor manipulated the objects on the toy in a continuous sequence. For each action, the actor's hand entered the screen near the goal object (thus immediately cueing which object was to be acted on), performed the action, and exited the screen again (3s). Between actions, the hand was off-screen and only the toy was visible (1s).

The timing of the actions was controlled such that the flow of movement appeared natural and continuous.

Figure 2. Example frames illustrating the predictive time window; arrows indicate the frames in which the agent's hand appears or disappears (Agent condition, top) and in which the spotlight focuses on the moving object (Ghost condition; bottom).



2.2.3 Ghost Condition

In the Ghost condition, the objects appeared to move on their own with a spotlight that focused on the current event (see Figure 2; bottom). The spotlight gradually illuminated each object before the object began to move and then faded again after the object finished moving. Between ghost events, there was always a period in which it was ambiguous where the spotlight would next begin to appear— and thus which object would subsequently move—to match the period in which the actor's hand was off-screen in the Agent condition. Like the hand, the spotlight also cued which object would move next. Neither the hand nor the spotlight moved across the screen between events but rather faded (spotlight) or left the screen (hand) in the same place. The spotlight was intended to match the saliency of the hand in the Agent condition as closely as possible, though there naturally remained perceptual differences between the two displays. Care was taken to ensure that its intensity and focus was equal for each object. Importantly, the statistical structure, order and timing of events were identical in both conditions.

2.3 Procedure

Toddlers were seated on their parent's lap throughout all phases of the experiment. First, toddlers were familiarized with the toy and allowed to explore it for one minute. The initial orientation of the toy relative to the child was counterbalanced between participants. Parents were asked not to influence their child's behavior throughout the testing session.

Caregivers and toddlers then moved to a chair approximately 63cm from the Tobii presentation screen. A nine-point calibration sequence was repeated until valid calibration data was acquired for at least eight calibration points. Following calibration, toddlers were shown the stimulus sequence until the presentation ended or the child became too fussy to continue the experiment. Caregivers were instructed to look aside during the calibration phase and not to influence their child throughout the session.

After the stimulus presentation, toddlers and caregivers returned to the table for a 'play' session in which the toddler was allowed to freely engage with the toy for two minutes or until they lost interest. The experimenter sat opposite them and pressed a hidden button to activate the light if the toddler performed the Effect pair, taking care to remain neutral. A camera facing the toddler recorded this session and behavior was later coded offline.

2.4 Eye Movement Data Processing

To ensure sufficient exposure to learn the sequences, we required individual fixation time across the entire experiment to be greater than one *SD* below the mean in each condition for inclusion in analysis. This resulted in the exclusion of 8 toddlers in the Agent condition and 10 in the Ghost condition (as noted in the *Participants* section). These participants yielded gaze data for less than 15% of the demonstration (mean looking times = 50.05*s*, *SD* = 34.10), corresponding to an average of 1.8 observations of each deterministic pair (i.e., less than one instance, on average, to demonstrate learning following the first exposure to the deterministic pair). Toddlers with looking times greater than one *SD* below the mean were included in analyses.

Eye movement data was exported from Tobii ClearView analysis software and separated into discrete fixations using a custom software program with a temporal filter of 100ms and a spatial filter of 30 pixels. Fixation data was imported into Matlab for further analysis. Regions of interest (ROI) were defined around each object (250 x 250 pixels), and a smaller ROI (130 x 130 pixels) was defined around the light due to its smaller size relative to the objects.

For the Agent condition, a fixation was considered predictive if it occurred in the time window from the moment the agent's hand appeared to perform the first action of a pair (entering the screen closest to the corresponding object and thus cueing the upcoming action) until the frame before it reappeared to perform the second action (Figure 2). This corresponds to the time in which the observer had enough information to predict the subsequent action before it actually occurred. For the Ghost condition, this time window was defined as from the first frame in which the spotlight began to highlight the first moving object (i.e., event) until the frame before it shifted to focus on the second moving object of a pair. Predictive time windows in both conditions were identical in length.

3.0 Results

To assess potential lower-level differences in visual attention between conditions, we first compared total looking times and number of predictive fixations made between conditions. Looking times did not differ between conditions, $t(56) = 1.47$, $p = .15$. Toddlers watched the stimulus videos for approximately 186.14 seconds ($SD = 54.22$) in the Agent condition and 165.01 seconds ($SD = 55.41$) in the Ghost condition. Toddlers in the two conditions did not differ in the average number of actions or events they watched, $t(57) = 1.11$, $p = .27$ ($M_{\text{Agent}} =$ 92.16 actions, $SD = 12.26$, $M_{\text{Ghost}} = 87.96$ events, $SD = 16.67$). There were no differences in the number of predictive fixations toddlers made throughout the experiment, $t(57) = 1.57$, $p = .12$ ($M_{\text{Agent}} = 307.65$, $SD = 103.08$, $M_{\text{Ghost}} = 266.36$, $SD = 98.71$), indicating that toddlers in one condition were not biased towards making more fixations during predictive time windows than in the other condition.

3.1 Learning over Trials

If toddlers learned the pair structure, we expected that predictive gaze fixations to the target object of each pair (labeled *Correct* predictions) would increase over trials as they repeatedly observed the deterministic relation between the two events of each pair. Proportions of correct fixations to target objects were calculated out of the total fixations to all ROIs, always excluding the object of the current action or event. Fixations to the light effect were considered correct predictions for the Effect pair and were included in the total fixations count (i.e., the denominator) for the Effect pair only (Eq. 1). Fixation proportions for the No-Effect pair never included looks to the light effect (Eq. 2).

$$Correct_{Object+light} = \frac{\#\, looks\ to\ target + \boldsymbol{light}}{total\ \#\, looks\ to\ all\ objects + \boldsymbol{light}} \quad (1)$$

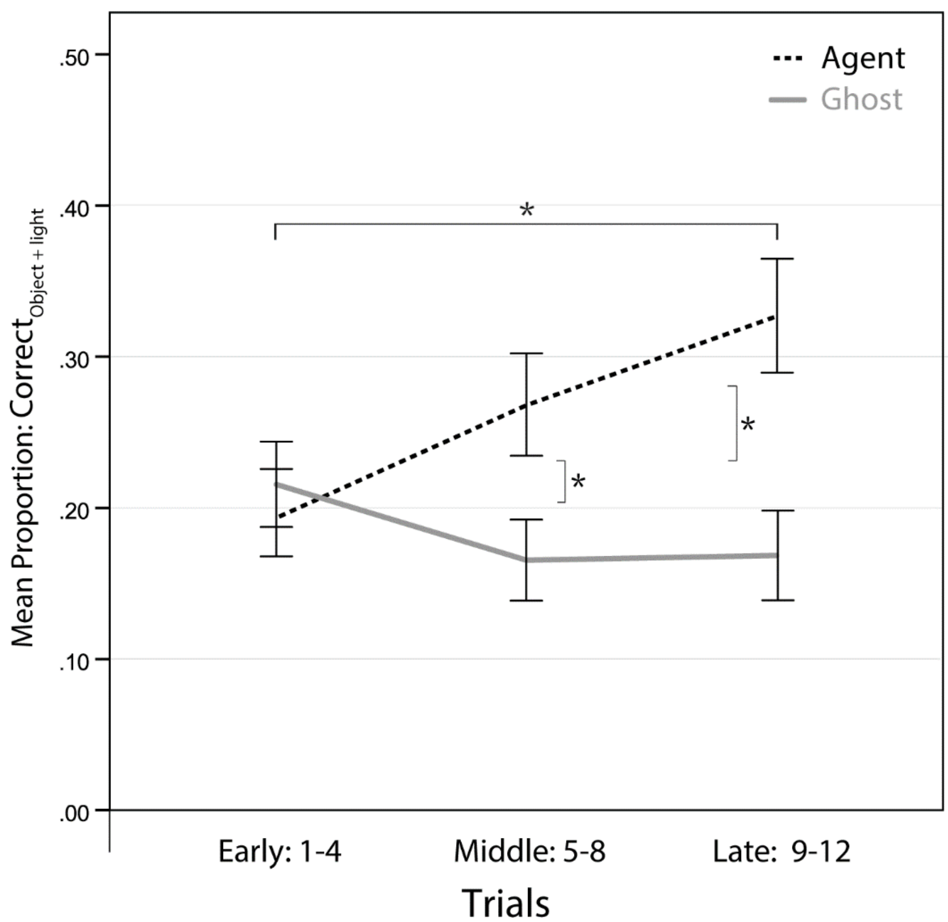
$$Correct_{Object} = \frac{\#\, looks\ to\ target}{total\ \#\, looks\ to\ all\ objects} \quad (2)$$

The 12 trials from each pair (Effect and No-effect) were collapsed into three time bins, with four trials in each bin. As an additional measure, we examined *Correct* predictions that excluded looks to the light effect for both pairs (Eq. 2) to assess whether the pattern was similar when only assessing fixations to the target objects.

Proportions of correct predictions for each pair (Eq. 1 for the Effect pair' Eq. 2 for the No-effect pair) were entered into a linear, model-based Generalized Estimating Equation (GEE) with an unstructured Working Correlation Matrix. GEEs are ideal to measure change over time for factors that are missing data for some entries (i.e., not all infants had data points for both pairs for each time bin; the number of missing data points did not differ between conditions in any time bin, *ps* > .37; Zeger, Liang, & Albert, 1988). Pair and Time Bin were entered as within-subjects repeated measures and Condition was a between-subjects measure. Findings revealed significant main effects of Condition, $\chi^2(1) = 22.40$, $p < .001$, Time Bin, $\chi^2(2) = 10.66$, $p = .005$, and Pair, $\chi^2 (1) = 16.99$, $p < .001$, a significant interaction between Condition and Time Bin, $\chi^2(2) = 11.52$, $p = .003$, and no other interactions (*p*s > .43).

Pairwise comparisons were conducted to probe the Condition by Time Bin interaction (Bonferroni-corrected critical $p = .008$, given the comparison between conditions for each time period [3] and comparisons between time points within each condition [2]). These revealed an upward learning curve in the Agent condition (Figure 3), such that the proportion of correct fixations increased from the first to the last bin (mean difference = .15 [*SEM* = .03], *p* < .001). No such increase in correct fixations occurred in the Ghost condition (*p*s > .81). The two conditions did not differ from one another in the first time bin (mean difference = -.01 [*SEM* = .03], *p* = .75). A significant difference emerged for the middle (mean difference = .10 [*SEM* = .03], *p* = .002) and last time bin (mean difference = .15 [*SEM* = .03], *p* < .001).

Figure 3. Toddlers' correct proportion scores (Eqs. 1-2) across pairs as a function of time. Bars represent standard errors.



A secondary analysis was conducted excluding looks to the light effect for the Effect pair (Eq. 2 for both pairs) which revealed a similar pattern: there were no differences between conditions during the first time bin (p = .36) but a significant difference emerged during the third time bin (p = .001).

3.2 Predicting target objects: correct vs. incorrect

If toddlers' learned to predict upcoming sequence events, we expected them to look more to target objects than to the remaining objects during the first action of each pair. Incorrect proportions were defined as the average number of fixations to the four remaining objects, excluding the object of the current or upcoming event, out of the total number of fixations (Eqs. 3 & 4). Comparing correct and incorrect proportions measured toddlers' preference for looking toward the correct target object, relative to other objects, before an object was acted upon.

$$Incorrect_{Object+light} = \frac{\#\ looks\ to\ other\ 4\ objects/4}{total\ \#\ looks\ to\ all\ objects+\boldsymbol{light}} \qquad (3)$$

$$Incorrect_{Object} = \frac{\#\ looks\ to\ other\ 4\ objects/4}{total\ \#\ looks\ to\ all\ objects} \qquad (4)$$

An ANOVA with Prediction (*Correct, Incorrect*) and Pair (*Effect*, *No-effect*) as within-subjects factors and Condition (*Agent*, *Ghost*) as a between-subjects factor revealed significant main effects of Pair, $F(1,54) = 8.08$, $p = .006$, = .13, and Condition, $F(1,54) = 5.95$, $p = .018$, $\eta_p^2 = .09$, and a significant Condition by Prediction interaction, $F(1,54) = 98.50$, $p = .018$, $\eta_p^2 = .10$. Planned comparisons to follow up on this interaction effect (Bonferroni-corrected critical *p*-value calculated as .025) revealed that *Correct* proportions were significantly greater than *Incorrect* proportions for toddlers in the *Agent* conditions (p = .003) whereas there was no difference for toddlers in the *Ghost* condition (p = .63; see Table 1; Figure 4, left).

As before, we performed secondary analyses in which we excluded the light effect in calculations for the Effect pair to more evenly match comparisons made between the two pairs

(Eq. 2&4 for both pairs). The same patterns emerged, with a significant interaction between Condition (*Agent*, *Ghost*) and Prediction ($Correct_{Object}$, $Incorrect_{Object}$), $F(1,54) = 9.00$, $p = .004$, $\eta_p^2 = .14$.

Table 1

Main dependent measures separated by condition

			Agent		Ghost	
DV	Pair	Factor	Mean	*SD*	Mean	*SD*
Correct vs. Incorrect	$Effect_{Object + Light}$	$Correct_{Object + Light}$	0.31*	*0.19*	0.24*	*0.14*
		$Incorrect_{Object + Light}$	0.17	*0.05*	0.19	*0.03*
	No-effect	Correct	0.22*	*0.22*	0.13*	*0.14*
		Incorrect	0.19	*0.05*	0.22	*0.03*
	$Effect_{Object}$	$Correct_{Object}$	0.20*	*0.16*	0.11*	*0.09*
		$Incorrect_{Object}$	0.20	*0.04*	0.22	*0.02*
Deterministic vs. Random	Effect	Deterministic trials	0.20*	*0.16*	0.11*	*0.09*
		Random trials	0.11	*0.12*	0.13	*0.11*
	No-effect	Deterministic trials	0.22*	*0.22*	0.13*	*0.14*
		Random trials	0.12	*0.10*	0.12	*0.16*
Action Performance	Effect	No. pairs performed	0.73	*0.87*	0.50	*0.61*
	No-effect	No. pairs performed	0.42	*0.64*	0.40	*0.68*
	Total	No. actions performed	25.19	*14.71*	21.65	*7.04*

Note. *$p < .025$ (difference between Agent and Ghost conditions). Fewer participants contributed action performance data, as some toddlers became too fussy to complete the play session following action observation.

3.3 Learning statistical context: deterministic vs. random transitions

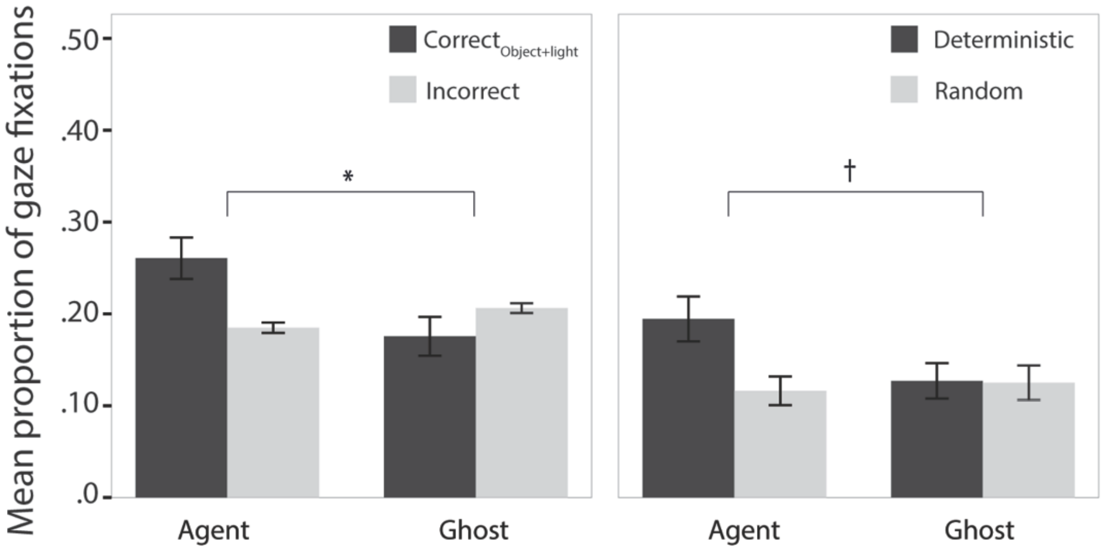
Targets occurred both within deterministic pairs and also at random outside of these pairs (see *Methods, 2.2.1*). For example, *open* occurred within a deterministic transition and was 100% predictable after *bend* had occurred. *Open* also occurred after other events, but featured low predictability (16.7%) at the onset of these (e.g., *slide*, *push*, *open*, or *spin*). We defined random transitions as the events that preceded the targets when they occurred randomly (Eq. 6) outside of their corresponding deterministic pair (i.e., a measure of 'chance', as defined in Tummeltshammer & Kirkham, 2013). We discarded trials in which the same action repeated itself (for example: *push* followed by *push*) as it was impossible to determine whether fixations during these trials were anticipations or the toddlers simply not moving their eyes. Analyzing the difference between deterministic (Eq. 5; same as *Correct*, Eq. 2) and random (Eq. 6) transitions enabled us to compare fixations to the same target objects in different statistical contexts. In these analyses, looks to the light effect were excluded entirely.

$$Deterministic = \frac{\#\, fixations\ to\ target\ (predictive\ trials)}{total\ \#\, fixations\ to\ objects} \qquad (5)$$

$$Random = \frac{\#\, fixations\ to\ target\ (non\text{−}predictive\ trials)}{total\ \#\, fixations\ to\ objects} \qquad (6)$$

An ANOVA with Transition (Deterministic, Random) and Pair (Effect, No-effect) as within-subject factors and Condition (Agent, Ghost) as a between-subjects factor revealed a marginally significant main effect of Transition, F(1,47) = 4.06, p = .05, η_p^2 = .08) and a marginally significant Condition by Transition interaction, F(1,47) = 3.65, p = .06, η_p^2 = .07 (Figure 4, right). Planned comparisons (Bonferroni-corrected critical p-value calculated as .025) revealed that more correct predictive fixations were made during the deterministic than random transitions in the Agent condition (p = .006; see Table 1). No such pattern was seen in the Ghost condition (p = .94).

Figure 4: Proportions of predictive fixations during action observation. *Left*: correct vs. incorrect (Eqs. 1-4, including the light in proportion scores for the *Effect* pair). *Right*: deterministic vs. random transitions (Eqs. 5-6). Bars represent standard errors. *p < .025, †p = .06



3.4 Behavior Coding

Behavior during the familiarization phase and during the first two minutes of the play session following the video session was coded for any instance in which the child made physical contact with one of the objects on the stimulus. A *success* was noted whenever the toddler made contact with an object and completed the action before removing his or her hand or visibly pausing (e.g., opening and closing the yellow door). An *attempt* was noted any time the child made contact with an object and completed part of the action before removing his or her hand or made an overt effort to manipulate an object in some way without success (e.g., opening the yellow door, but leaving it open). If the child only placed his or her hand on an object without attending to it or attempting to manipulate it (i.e. no attempt to perform an 'action'), these instances were considered invalid and were excluded from analyses. We counted the number of Effect and No-effect pairs that each toddler reproduced and the total number of actions they produced during the familiarization and post-observation phases. Fifty percent of the total participants were coded by a second coder. There was strong agreement between the two coders, Cohen's $\kappa = .729$, $p < .001$.

During the post-observation play session, 51% of the toddlers in this study performed at least one of the deterministic action pairs (Table 1). Toddlers in the Agent condition were not simply more active than in the Ghost condition overall: the total number of actions performed (the sum of successes and attempts) did not differ between conditions, $t(44) = .99$, $p = .33$. There were no differences between conditions in the frequency of performing either pair (Effect pair: $t(44) = 1.01$, $p = .32$; No-effect pair: $t(44) = .12$, $p = .91$). Across conditions, there were no differences in the number of times toddlers produced the Effect and No-effect pairs, $t(45) = 1.40$, $p = .17$, $r = -.09$.

Correlations between the proportions of correct fixations and the number of action pairs that toddlers produced were conducted separately for Effect and No-effect pairs. For the Effect pair, correct predictive fixations (Eq. 1) were correlated with action performance in the Agent condition ($r = .37$, N = 26, $p = .03$, one-tailed). This relation was not significant for toddlers in the Ghost condition ($r = .17$, N = 20, $p = .20$). There was no significant difference between these two correlations, however (Fisher r-to-z transformation, $Z = .68$, $p = .25$; Kenny, 1987). For the No-effect pair, there were no significant correlations between gaze behavior (Eq. 3) and action performance in either condition (*p*s > .30).

To assess the possibility that toddlers who were better at detecting the sequence regularities were simply more active to begin with, we assessed toddlers' baseline activity during the familiarization phase prior to action observation. Baseline activity (defined as the mean number of actions performed per minute) did not correlate toddlers' proportions of correct fixations for either the Effect pair (including or excluding the effect) or the No-effect pair (*ps* > .62).

4.0 Discussion

In a novel eye-tracking paradigm, we investigated the role of statistical learning in action prediction and performance in young toddlers. Previous research has shown that statistical learning extends into the action domain and facilitates segmentation of actions into meaningful subunits. Here, we examined whether toddlers correctly predicted upcoming events of an action or ghost event sequence via observation of deterministic transitional probabilities between sequential events.

We evaluated three related measures of statistical learning that targeted different underlying questions regarding toddlers' improvement over trials and their overall predictive looking performance. All three measures demonstrated a similar pattern of results, with toddlers in the Agent condition demonstrating learning of the statistical structure in the sequence whereas those in the Ghost condition did not. Specifically, we found that only toddlers in the Agent condition showed a linear increase in their correct predictions for deterministic pairs over the course of the experiment. These toddlers also anticipated correct locations more frequently than incorrect locations during deterministic pairs, and made a greater rate of anticipatory fixations to target objects during deterministic trials than random trials. Toddlers in the Ghost condition, in contrast, did not differentiate between correct and incorrect locations and did not look to target objects more during deterministic transitions than during random transitions. Finally, for toddlers in the Agent condition, rates of correct predictions also correlated with their own spontaneous action performance following action observation. Our findings thus provide new evidence that on-line statistical learning directly contributes to predictive gaze behaviors during toddlers' observation of continuous action sequences.

4.1 Learning from observing an agent's actions vs. visual events

Toddlers who observed an actor were more accurate in predicting actions compared to those who observed a ghost sequence, despite being exposed to identical sequential information and demonstrating similar visual attention. This finding complements previous developmental work concerning the relative importance of agentive cues when young children are initially forming representations about objects and events in their environments (Hopper, Flynn, Wood, & Whiten, 2010). Hopper and colleagues showed that, for preschool children, learning to use a complex tool via observation was less effective in a ghost condition during which a tool's function was observed without any human action, compared with a condition in which a person demonstrated the actions with a tool (Hopper et al., 2010). These authors posited that, when a task is novel and challenging, young children need to watch an agent performing all the steps in an action sequence in order to imitate the entire behavior. Our results extend this notion by showing that observing an actor facilitates toddlers' abilities to use their statistical learning skills for action prediction.

Social cues help infants notice statistical properties of objects in noisy visual scenes (Wu, Gopnik, Richardson, & Kirkham, 2010). With limited attention and cognitive resources, social cues likely serve as an adaptive filter to help toddlers and infants select relevant information from the environment. Although general attention to the displays did not differ between conditions, it is possible that the presence of a social agent, as depicted by the hand, selectively directed anticipatory attention toward the sequential patterns in the ongoing actions.

The reason why cues such as hands and faces selectively guide infants' attention is an open question. In our paradigm, there were no pedagogical, facial, or other bodily cues present; observation of the actor's hand alone drove the effect observed between conditions. One possible explanation for this effect echoes current embodied accounts of the unique role of the motor system during action observation (Hunnius & Bekkering, 2014; Woodward & Gerson, 2014). Infants become better at predicting actions as they acquire motor experience with the observed movements, suggesting that motor activity during observation facilitates their predictions (Kanakogi & Itakura, 2011; Stapel et al., 2016). Compelling evidence from adult research has shown that motor activation is causally related to action prediction (Elsner et al., 2013) and similar evidence has been found in infancy (Marshall, Young, & Meltzoff, 2011; Stapel et al., 2016; van Elk, van Schie, Hunnius, Vesper, & Bekkering, 2008). Our findings are consistent with the idea that the motor system facilitates action predictions during observation. Though we did not directly test motor activity here, it is a theoretically grounded possibility that we are currently pursuing in follow-up studies.

An alternative explanation for the different patterns found in the Agent versus Ghost conditions relies on low-level perceptual accounts. For example, the toddlers need not represent anything about actions or agents if the saliency of the events differed between conditions as a function of the use of a hand versus a spotlight. Although we accounted for looking time and number of fixations, it is possible that some unidentifiable difference in saliency existed between conditions. If this were the case, however, we would not expect a relation between action prediction and reproduction specifically for the Agent condition.

4.2 Learning in infancy: structure-detecting skills

Several learning mechanisms have been identified that may contribute to cognitive development in complementary ways. Researchers have, for instance, provided empirical and conceptual evidence for accounts of causal (e.g., Gopnik, Sobel, Schulz, & Glymour, 2001), associative (e.g., Ray & Heyes, 2011), and statistical learning (e.g., Kirkham et al., 2002). In our experiment, we referred to the learning mechanism taking place as *statistical* learning, mainly because the information that we manipulated in our experiment was statistical in nature. We consider these learning mechanisms to be components of the same set of general structure-detecting skills. Our focus here was to provide evidence for how prediction can arise from the statistics contained in action sequences, and in particular, how action processing may differ from general visual processing.

Despite providing evidence for learning, we found that toddlers' proportion scores were generally low and did not exceed 30% except in the Agent condition on later trials. One reason for this is likely due to the fact that, unlike previous studies, we only examined toddlers' anticipatory (rather than reactive) looks to upcoming targets. In a recent experiment investigating spatiotemporal learning in 8-month-olds, approximately 35% of infants' gaze fixations were anticipatory, whereas the remaining 65% were reactive (Tummeltshammer & Kirkham, 2013). For more complex stimuli (i.e., multiple objects presented simultaneously to which infants must distribute their visual attention), up to 95% of infants' fixations were reactive rather than predictive (Tummeltshammer, Pomiechowska, Gliga & Kirkham, 2015). In the current study, we introduced the additional demands of tracking statistical relations between actions and events without spatial information to help guide predictions. We found that toddlers can make anticipatory eye movements for these complex sequences when they observed the Agent condition, but not the Ghost condition.

4.3 The influence of action-effects on learning

Interestingly, our findings revealed no significant interactions between toddlers' correct anticipations and the presence of the action-effect (i.e., Effect vs. No-effect pairs), and the pattern of results were similar when the action-effect was removed from analyses. This suggests that the evidence for learning statistical regularities did not depend on whether the paired actions caused an effect. Rather, those toddlers who learned the associations between observed actions did not need to rely on other supportive cues (consistent with Waismeyer et al., 2015). The light effect's distal relation to its associated action may have even hindered learning the action-effect association, although the finding that more toddlers selectively reproduced the Effect pair speaks against this possibility.

A separate question raised by this research is whether the transitions between the actions and their effects are encoded qualitatively differently from the transitions between actions. The most immediate sensory consequence of an action or an event is the change in the environment directly caused by the action itself, such as the opening of the door in our experiment. These could be considered proximal action-effects, and the light effect a more distal effect or a separate goal-related outcome altogether. We did not specifically address these questions in our analyses, but they represent an interesting direction to pursue in follow-up studies.

4.4 Action performance

The relationship between how toddlers perceive actions and their own emerging action capabilities is an important developmental phenomenon. Our findings indicate a relation between uninstructed learning from observation and spontaneous action performance that was not due to a priori differences in toddlers' motor behavior. For the toddlers who observed a human actor, those who were better at predicting the second action of the action-effect pair were also more likely to integrate what they learned into their own behavior by performing this pair more frequently themselves. Given the intentional lack of any instructional or pedagogical cue in the current research, it is remarkable that there was an immediate transfer between observation and reproduction on such a brief timescale. Some insight into the underlying mechanism comes from the reinforcement learning literature, which has shown that transitional probabilities can be learned in the absence of any outcomes, but only manifests in behavior when action choices lead to a reward (Gläscher, Daw, Dayan, & O'Doherty, 2010). Possibly, the action-effect was interpreted as a rewarding outcome that enabled or motivated some toddlers to reproduce it. There was no correlation between predictions and action performance for the pair that did not cause an effect, nor when toddlers did not observe a human actor.

For the toddlers who did not reproduce any action pairs, exploration may have simply been more interesting than activating the light or imitating the observed sequence. Exploration is a critical developmental behavior that allows discovery of new environmental features (Hopper et al., 2010). Though important in certain contexts, adult instruction causes children to limit their exploratory play in order to focus on performing target actions (Bonawitz, Shafto, Gweon, Goodman, Spelke, & Schulz, 2011). The toddlers who did reproduce effect pairs may have interpreted the agent's demonstration as an instructional cue, and constrained their behavior to those actions more likely to cause a salient outcome (e.g., Bushnell & Boudreau, 1998).

4.5 Conclusion

Our findings provide evidence that toddlers use their statistical learning skills to learn the structure in observed action sequences and make predictions about upcoming actions. Making accurate predictions depended on whether they observed an agent or a ghost sequence, providing further support for the unique role of observing human agents on toddlers' action processing. Further, we found evidence for a correlation between observational learning and action production, which was constrained by both action cues and action effects. In sum, our study shows that statistical regularities in observed action sequences are a powerful learning opportunity for toddlers and support the developing ability to generate predictions about future actions. From observing their parents preparing a meal and knowing what the outcome will be, to acquiring the ability to perform increasingly complex action sequences themselves, young children make use of diverse statistical and social information in order to understand and predict the structure underlying human actions.

References

Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: Active inference in the motor system. *Brain Structure and Function, 218*(3), 611-643. doi: 10.1007/s00429-012-0475-5

Amso, D., & Davidow, J. (2012). The development of implicit learning from infancy to adulthood: item frequencies, relations, and cognitive flexibility. *Developmental Psychobiology*, *54*(6), 664-673. doi: 10.1002/dev.20587

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development, 72*(3), 708-717. doi: 10.1111/1467-8624.00310

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, *120*(3), 322–330. doi: 10.1016/j.cognition.2010.10.001

Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. Cognition, 120(3), 331-340. doi: 10.1016/j.cognition.2010.12.001

Bushnell, E. W., & Boudreau, J. P. (1998). Exploring and exploiting objects with the hands during infancy. In Connolly, K. J. (Ed.), *The Psychobiology of the Hand. Clinics in Developmental Medicine*, *147*, 144-161. London, England: Mac Keith Press.

Canfield, R. L., & Haith, M. M. (1991). Young infants' visual expectations for symmetric and asymmetric stimulus sequences. *Developmental Psychology, 27*(2), 198. doi: 10.1037/0012-1649.27.2.198

Danish, D., & Russell, J. (2007). The role of ‘action-effects’ and agency in toddlers’ imitation. *Cognitive Development, 22*(1), 69-76. doi: 10.1016/j.cogdev.2006.08.001

Elsner, B., Hommel, B., Mentschel, C., Drzezga, A., Prinz, W., Conrad, B., & Siebner, H. (2002). Linking actions and their perceivable consequences in the human brain. NeuroImage, 17(1), 364-372. doi: 10.1006/nimg.2002.1162

Elsner, C., D'Ausilio, A., Gredebäck, G., Falck-Ytter, T., & Fadiga, L. (2013). The motor cortex is causally related to predictive eye movements during action observation. *Neuropsychologia, 51*(3), 488-492. doi: 10.1016/j.neuropsychologia.2012.12.007

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience, 9*(7), 878-879. doi: 10.1038/nn1729

Fawcett, T. W., Skinner, A. M. J., & Goldsmith A. R. (2002). A test of imitative learning in starling using two-action method with an enhanced ghost control. *Animal Behaviour*, *64*, 547–556. doi: 10.1006/anbe.2002.3092

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, *424*(6950), 769-771. doi: 10.1038/nature01861

Gerson, S. A., & Woodward, A. L. (2013). The goal trumps the means: Highlighting goals is more beneficial than highlighting means in means-end training. *Infancy, 18*(2), 289-302. doi: 10.1111/j.1532-7078.2012.00112.x

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron, 66*(4), 585-595. doi: 10.1016/j.neuron.2010.04.016

Gredebäck, G., & Kochukhova, O. (2009). Goal anticipation during action observation is influenced by synonymous action capabilities, a puzzling developmental study. *Experimental brain research, 202*(2), 493-497. doi: 10.1007/s00221-009-2138-1

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology, 37*(5), 620-629. doi: 10.1037/0012-1649.37.5.620

Hopper, L. M. (2010). 'Ghost' experiments and the dissection of social learning in humans and animals. *Biological Reviews, 85*(4), 685-701. doi: 10.1111/j.1469-185X.2010.00120.x

Hopper, L. M., Flynn, E. G., Wood, L. A. N., & Whiten, A. (2010). Observational learning of tool use in children: Investigating cultural spread through diffusion chains and learning mechanisms through ghost displays. *Journal of Experimental Child Psychology, 106*(1), 82-97. doi: 10.1016/j.jecp.2009.12.001

Hopper, L. M., Lambeth, S. P., Schapiro, S. J., & Whiten, A. (2014). The importance of witnessed agency in chimpanzee social learning of tool use. *Behavioural Processes, 112*, 120-129. doi: 10.1016/j.beproc.2014.10.009

Hunnius, S., & Bekkering, H. (2010). The early development of object knowledge: A study of infants' visual anticipations during action observation. *Developmental Psychology, 46*(2), 446-454. doi: 10.1037/a0016543

Hunnius, S., & Bekkering, H. (2014). What are you doing? How active and observational experience shape infants' action understanding. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1644), 20130490. doi: 10.1098/rstb.2013.0490

Jovanovic, B., Kiraly, I., Elsner, B., Gergely, G., Prinz, W., & Aschersleben, G. (2007). The role of effects for infants' perception of action goals. *Psychologia 50*(4), 273-290. doi: 10.2117/psysoc.2007.273

Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nature Communications, 2*, 341. doi: 10.1038/ncomms1342

Kenny, David A. (1987*). Statistics for the social and behavioral sciences*. Little, Brown.

Keysers, C., Kaas, J. H., & Gazzola, V. (2010). Somatosensation in social perception. *Nature Reviews Neuroscience, 11*(6), 417-428. doi: 10.1038/nrn2919

Kirkham, N., Slemmer, J., & Johnson, S. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35-B42. doi:10.1016/S0010-0277(02)00004-5

Klein, A. M., Hauf, P., & Aschersleben, G. (2006). The role of action effects in 12-month-olds' action control: A comparison of televised model and live model. *Infant Behavior and Development, 29*(4), 535-544. doi: 10.1016/j.infbeh.2006.07.001

Kochukhova, O., & Gredebäck, G. (2010). Preverbal infants anticipate that food will be brought to the mouth: An eye tracking study of manual feeding and flying spoons. *Child Development, 81*(6), 1729-1738. doi: 10.1111/j.1467-8624.2010.01506.x

Krogh, L., Vlach, H., & Johnson, S. P. (2013). Statistical learning across development: Flexible yet constrained. *Frontiers in Psychology*, *3*, 598. doi: 10.3389/fpsyg.2012.00598

Marshall, P. J., Young, T., & Meltzoff, A. N. (2011). Neural correlates of action observation and execution in 14-month-old infants: An event-related EEG desynchronization study. *Developmental Science, 14*(3), 474-480. doi: 10.1111/j.1467-7687.2010.00991.x

Paulus, M., Hunnius, S., van Wijngaarden, C., Vrins, S., van Rooij, I., & Bekkering, H. (2011). The role of frequency information and teleological reasoning in infants' and adults' action prediction. *Developmental Psychology, 47*(4), 976-983. doi: 10.1037/a002

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*(2), 129-154. doi:10.1080/713752551

Ray, E., & Heyes, C. (2011). Imitation in infancy: the wealth of the stimulus. *Developmental Science*, *14*(1), 92-105. doi: 10.1111/j.1467-7687.2010.00961.

Reid, V. M., Hoehl, S., Grigutsch, M., Groendahl, A., Parise, E., & Striano, T. (2009). The neural correlates of infant and adult goal prediction: Evidence for semantic processing systems. *Developmental Psychology, 45*(3), 620-629. doi: 10.1037/a0015209

Romberg, A. R., & Saffran, J. R. (2012). Expectancy learning from probabilistic input by infants. *Frontiers in Psychology, 3*, 610. doi: 10.3389/fpsyg.2012.00610

Ruffman, T., Taumoepeau, M., & Perkins, C. (2012). Statistical learning as a basis for social understanding in children. *British Journal of Developmental Psychology, 30*(1), 87-104. doi: 10.1111/j.2044-835X.2011.02045.x

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926 - 1928. doi: 10.1126/science.274.5294.1926

Saylor, M. M., Baldwin, D. A., Baird, J. A., & LaBounty, J. (2007). Infants' on-line segmentation of dynamic human action. *Journal of Cognition and Development, 8*(1), 113-128. doi: 10.1080/15248370709336996

Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences*, *11*(5), 211-218.

Stahl, A. E., Romberg, A. R., Roseberry, S., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Infants segment continuous events using transitional probabilities. *Child Development* 85(5), 1821-1826. doi: 10.1111/cdev.12247

Stapel, J. C., Hunnius, S., van Elk, M., & Bekkering, H. (2010). Motor activation during observation of unusual versus ordinary actions in infancy. *Social Neuroscience, 5*(5-6), 451-460. doi: 10.1080/17470919.2010.490667

Stapel, J. C., Hunnius, S., Meyer, M., & Bekkering, H. (2016). Motor system contribution to action prediction: Temporal accuracy depends on motor experience. *Cognition, 148*, 71-78. doi: 10.1016/j.cognition.2015.12.007

Thompson, D. E., & Russell, J. (2004). The ghost condition: Imitation versus emulation in young children's observational learning. *Developmental Psychology, 40*(5), 882. doi: 10.1037/0012-1649.40.5.882

Tummeltshammer, K. S., & Kirkham, N. Z. (2013). Learning to look: Probabilistic variation and noise guide infants' eye movements. *Developmental Science, 16*(5), 760-771. doi: 10.1111/desc.12064

Tummeltshammer, K., Pomiechowska, B., Gliga, T. & Kirkham, N. (2015). *Spatiotemporal regularity and infants' visual predictions: an ERP study.* Poster presented at the International Conference on Interdisciplinary Advances in Statistical Learning, San Sebastian, Spain.

Van Elk, M., van Schie, H. T., Hunnius, S., Vesper, C., & Bekkering, H. (2008). You'll never crawl alone: Neurophysiological evidence for experience-dependent motor resonance in infancy. *NeuroImage, 43*(4), 808-814. doi: 10.1016/j.neuroimage.2008.07.057

Waismeyer, A., Meltzoff, A. N., & Gopnik, A. (2015). Causal learning from probabilistic events in 24-month-olds: An action measure. *Developmental Science, 18*(1), 175-182. doi: 10.1111/desc.12208

Wentworth, N., Haith, M., & Hood, R. (2002). Spatiotemporal regularity and inter-event contingencies as information for infants' visual expectations. *Infancy*, *3*(3), 303-321. doi: 10.1207/S15327078IN0303_2

Woodward, A. L., & Gerson, S. A. (2014). Mirroring and the development of action understanding. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1644), 20130181. doi: 10.1098/rstb.2013.0181.

Woodward, A. L., & Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science, 11*(1), 73-77. doi: 10.1111/1467-9280.00218

Wu, Rachel and Gopnik, A. and Richardson, D.C. and Kirkham, Natasha Z. (2010) Social cues support learning about objects from statistics in infancy. In Ohlsson, S. and Catrambone, R. (Eds.) *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (1228-1233). Austin, Texas, USA: Cognitive Science Society.

Zeger, S. L., Liang, K.-Y., & Albert, P. S. (1988). Models for longitudinal data: A generalized estimating equation approach. *Biometrics, 44*(4), 1049-1060. doi: 10.2307/2531734