

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:<https://orca.cardiff.ac.uk/id/eprint/96942/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Monroy, Claire, Gerson, Sarah and Hunnius, Sabine 2017. Toddlers' action prediction: statistical learning of continuous action sequences. *Journal of Experimental Child Psychology* 157 , pp. 14-28. 10.1016/j.jecp.2016.12.004

Publishers page: <http://dx.doi.org/10.1016/j.jecp.2016.12.004>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38

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

Claire D. Monroy <sup>a,b</sup>, Sarah A. Gerson <sup>a,c</sup>, Sabine Hunnius <sup>a</sup>

<sup>a</sup> Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen

<sup>b</sup> Ohio State University Wexner Medical Center

<sup>c</sup> Cardiff University, School of Psychology

*In Press, Journal of Experimental Child Psychology, 2017*

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22

## 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, Vrans, 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

1 initially investigated in the auditory domain, and recently researchers have shown that  
2 statistical learning also allows infants to extract structure from visual input as well (Kirkham,  
3 Slemmer & Johnson, 2002). The present study aimed to determine whether toddlers can  
4 discover the statistical regularities in action sequences and accurately predict upcoming actions  
5 and their effects during online observation.

### 6 1.1 Statistical learning: sensitivity to sequential structure

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

18       For example, in a study by Baldwin and colleagues (2001), 10-month-olds were  
19 familiarized to videos of daily actions such as sweeping. In a subsequent test phase, they  
20 watched versions of these videos in which pauses were inserted in the midst of the actor's  
21 motion ('interrupting' videos), or at natural breakpoints in the action stream ('intact' videos).  
22 Infants increased visual attention to the interrupting videos, which was interpreted as evidence

1 that they perceived a violation in the expected structure. Additional research has shown that,  
2 rather than being restricted to highly familiar events, infants can also learn the structure of  
3 novel, abstract sequences (e.g., the dancing of a starfish) in which the only cues for  
4 segmentation are the TPs between dynamic events (Stahl, Romberg, Roseberry, Golinkoff, &  
5 Hirsh-Pasek, 2014).

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

1 perceiving human action engages unique cognitive mechanisms (Schubotz, 2007) that may  
2 extend to how statistical regularities in action sequences are learned, as we discuss in the  
3 following section.

#### 4 1.2 Relations between action prediction and execution in development

5 A growing body of evidence suggests that action prediction is constrained by infants'  
6 developing action repertoire. The ability to predict action outcomes relates to developing  
7 motor capabilities or experiences (e.g., Gredebäck & Kochukhova, 2009; Kanakogi & Itakura,  
8 2011; Kochukhova & Gredebäck, 2010; Stapel, Hunnius, Meyer, & Bekkering, 2016). These  
9 developmental findings are consistent with embodied accounts of action observation, which  
10 propose that action prediction arises from activated representations in the observer's own  
11 motor system (Adams, Shipp, & Friston, 2013; Elsner, D'Ausilio, Gredebäck, Falck-Ytter, &  
12 Fadiga, 2013; Flanagan & Johansson, 2003; Hunnius & Bekkering, 2014; Keyzers, Kaas, &  
13 Gazzola, 2010). According to these embodied accounts, observing actions engages the  
14 observer's own motor system and provides an additive learning benefit beyond other forms of  
15 observational learning.

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

1 display method to examine whether action observation uniquely facilitates prediction of  
2 sequential actions.

### 3 1.3 Actions and their effects

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

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



1 action reproduction following implicit learning as an additional exploratory measure in the  
2 current experiment.

### 3 1.4 The current study

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

## 14 2.0 Methods

### 15 2.1 Participants

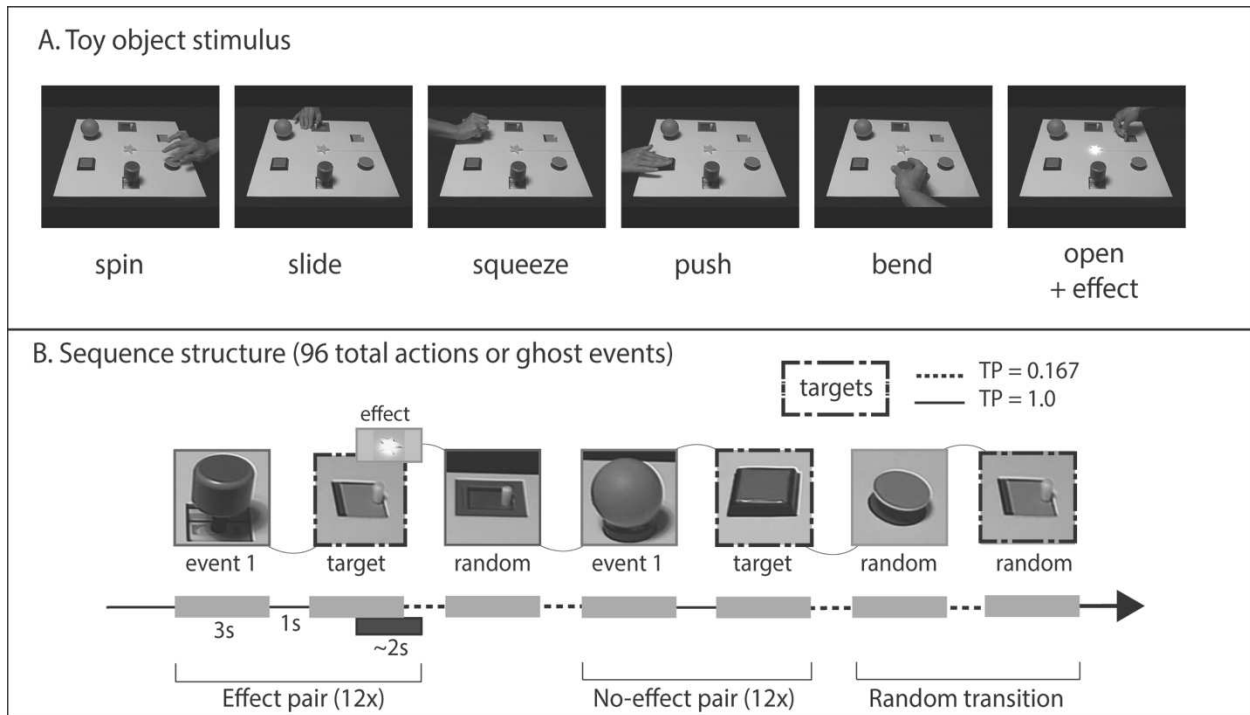
16 Eighty-six toddlers participated in this study. Twenty-seven toddlers were excluded due to  
17 insufficient gaze data (see Analysis section), excessive fussiness or failure of the eye-tracker.  
18 Fifty-nine toddlers remained in the final sample (mean age = 19.1 months, range: 18.5-20.5  
19 months,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

## 1 2.2 Stimuli

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

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

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



1

2

### 2.2.1 Sequence

3

We created four pseudo-randomized sequences of 96 individual events using the program

4

Mix (van Casteren & Davis, 2006). These sequences contained two deterministic pairs (TP =

5

1.0). The remaining sequence events occurred randomly (TP = .167 or 1/6; Figure 1B). For

6

clarity, we will refer to the second events of the deterministic pairs as *targets*, as these were

7

the events that became predictable as the sequence unfolded. One deterministic pair caused a

8

central star on the toy to light up (Effect pair), and the second pair had no effect (No-effect

9

pair). The effect occurred at a natural mid-point of the target action or event. For example,

10

during *open*, the light turned on the moment the yellow door was fully open and turned off

11

after it closed. The objects corresponding to Effect and No-effect pairs were counterbalanced

12

across toddlers.

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

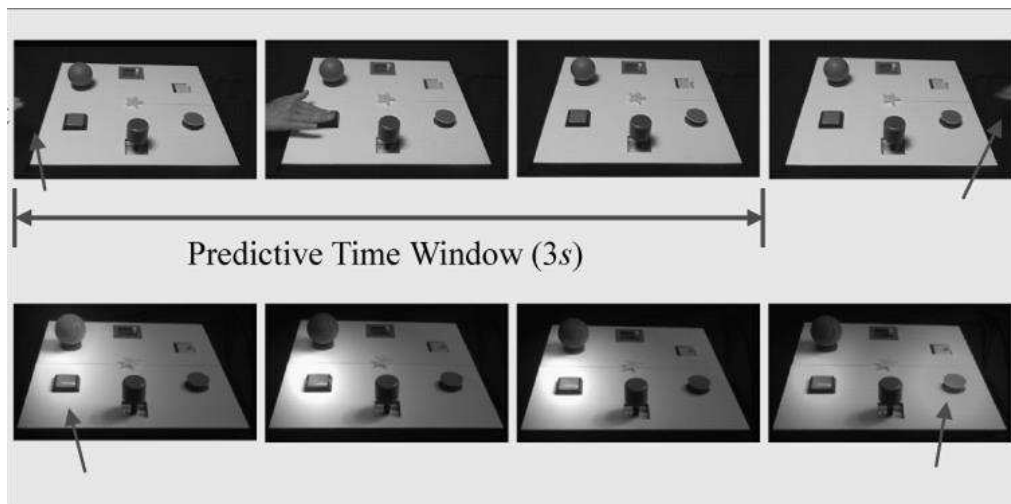
10           Sequences were divided into four blocks, with the viewing orientation rotated during  
11 each block to ensure that the toddler could not predict the next action or event from its  
12 location on the screen. Each block lasted approximately 90s and consisted of 24 events. Brief  
13 animations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

### 17           2.2.2 Agent Condition

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

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

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



### 6 2.2.3 Ghost Condition

7 In the Ghost condition, the objects appeared to move on their own with a spotlight that  
 8 focused on the current event (see Figure 2; bottom). The spotlight gradually illuminated each  
 9 object before the object began to move and then faded again after the object finished moving.  
 10 Between ghost events, there was always a period in which it was ambiguous where the  
 11 spotlight would next begin to appear— and thus which object would subsequently move—to  
 12 match the period in which the actor's hand was off-screen in the Agent condition. Like the  
 13 hand, the spotlight also cued which object would move next. Neither the hand nor the spotlight  
 14 moved across the screen between events but rather faded (spotlight) or left the screen (hand)  
 15 in the same place. The spotlight was intended to match the saliency of the hand in the Agent

1 condition as closely as possible, though there naturally remained perceptual differences  
2 between the two displays. Care was taken to ensure that its intensity and focus was equal for  
3 each object. Importantly, the statistical structure, order and timing of events were identical in  
4 both conditions.

### 5 2.3 Procedure

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

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

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

### 21 2.4 Eye Movement Data Processing

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

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

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

1 frame before it shifted to focus on the second moving object of a pair. Predictive time windows  
2 in both conditions were identical in length.

### 3 3.0 Results

4 To assess potential lower-level differences in visual attention between conditions, we first  
5 compared total looking times and number of predictive fixations made between conditions.  
6 Looking times did not differ between conditions,  $t(56) = 1.47, p = .15$ . Toddlers watched the  
7 stimulus videos for approximately 186.14 seconds ( $SD = 54.22$ ) in the Agent condition and  
8 165.01 seconds ( $SD = 55.41$ ) in the Ghost condition. Toddlers in the two conditions did not  
9 differ in the average number of actions or events they watched,  $t(57) = 1.11, p = .27$  ( $M_{\text{Agent}} =$   
10  $92.16$  actions,  $SD = 12.26$ ,  $M_{\text{Ghost}} = 87.96$  events,  $SD = 16.67$ ). There were no differences in the  
11 number of predictive fixations toddlers made throughout the experiment,  $t(57) = 1.57, p = .12$   
12 ( $M_{\text{Agent}} = 307.65, SD = 103.08, M_{\text{Ghost}} = 266.36, SD = 98.71$ ), indicating that toddlers in one  
13 condition were not biased towards making more fixations during predictive time windows than  
14 in the other condition.

#### 15 3.1 Learning over Trials

16 If toddlers learned the pair structure, we expected that predictive gaze fixations to the  
17 target object of each pair (labeled *Correct* predictions) would increase over trials as they  
18 repeatedly observed the deterministic relation between the two events of each pair.  
19 Proportions of correct fixations to target objects were calculated out of the total fixations to all  
20 ROIs, always excluding the object of the current action or event. Fixations to the light effect  
21 were considered correct predictions for the Effect pair and were included in the total fixations



1 count (i.e., the denominator) for the Effect pair only (Eq. 1). Fixation proportions for the No-  
 2 Effect pair never included looks to the light effect (Eq. 2).

$$3 \quad \text{Correct}_{\text{Object+light}} = \frac{\# \text{ looks to target+light}}{\text{total \# looks to all objects+light}} \quad (1)$$

$$4 \quad \text{Correct}_{\text{Object}} = \frac{\# \text{ looks to target}}{\text{total \# looks to all objects}} \quad (2)$$

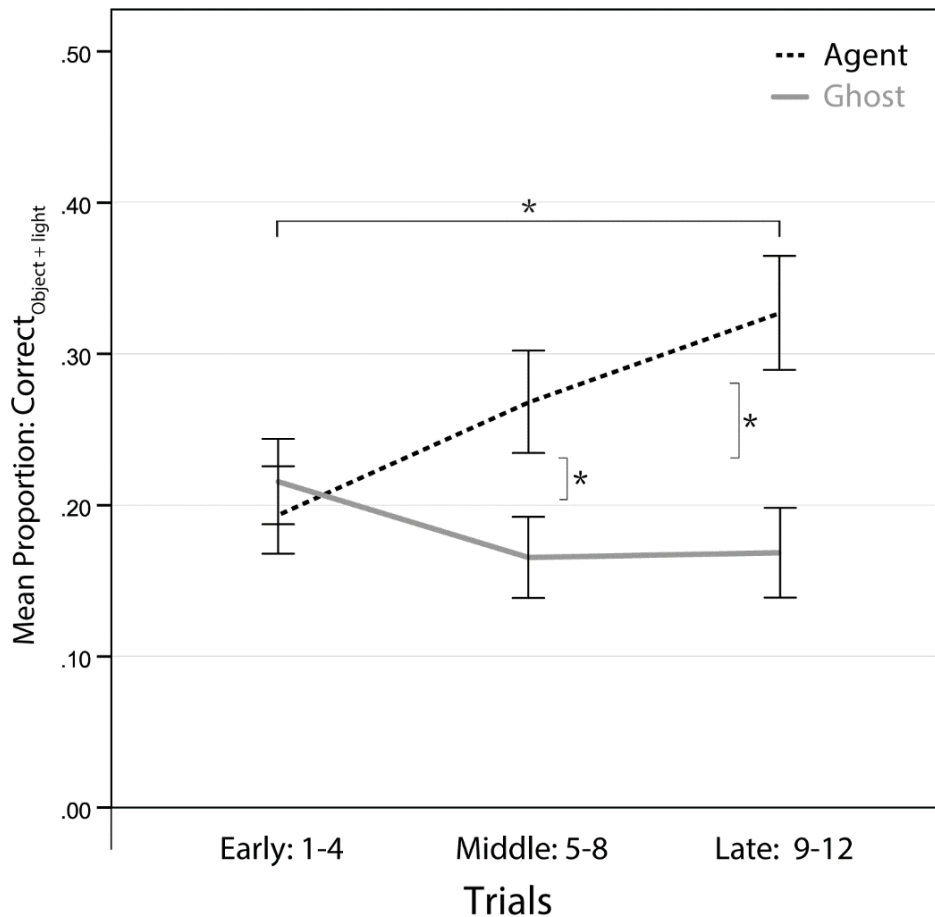
5 The 12 trials from each pair (Effect and No-effect) were collapsed into three time bins, with four  
 6 trials in each bin. As an additional measure, we examined *Correct* predictions that excluded  
 7 looks to the light effect for both pairs (Eq. 2) to assess whether the pattern was similar when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

9 Proportions of correct predictions for each pair (Eq. 1 for the Effect pair' Eq. 2 for the No-  
 10 effect pair) were entered into a linear, model-based Generalized Estimating Equation (GEE) with  
 11 an unstructured Working Correlation Matrix. GEEs are ideal to measure change over time for  
 12 factors that are missing data for some entries (i.e., not all infants had data points for both pairs  
 13 for each time bin; the number of missing data points did not differ between conditions in any  
 14 time bin,  $ps > .37$ ; Zeger, Liang, & Albert, 1988). Pair and Time Bin were entered as within-  
 15 subjects repeated measures and Condition was a between-subjects measure. Findings revealed  
 16 significant main effects of Condition,  $\chi^2(1) = 22.40, p < .001$ , Time Bin,  $\chi^2(2) = 10.66, p = .005$ ,  
 17 and Pair,  $\chi^2(1) = 16.99, p < .001$ , a significant interaction between Condition and Time Bin,  $\chi^2(2)$   
 18  $= 11.52, p = .003$ , and no other interactions ( $ps > .43$ ).

19 Pairwise comparisons were conducted to probe the Condition by Time Bin interaction  
 20 (Bonferroni-corrected critical  $p = .008$ , given the comparison between conditions for each time  
 21 period [3] and comparisons between time points within each condition [2]). These revealed an  
 22 upward learning curve in the Agent condition (Figure 3), such that the proportion of correct

1 fixations increased from the first to the last bin (mean difference = .15 [ $SEM = .03$ ],  $p < .001$ ).  
 2 No such increase in correct fixations occurred in the Ghost condition ( $ps > .81$ ). The two  
 3 conditions did not differ from one another in the first time bin (mean difference = -.01 [ $SEM$   
 4 = .03],  $p = .75$ ). A significant difference emerged for the middle (mean difference = .10 [ $SEM$   
 5 = .03],  $p = .002$ ) and last time bin (mean difference = .15 [ $SEM = .03$ ],  $p < .001$ ).

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



8

9 A secondary analysis was conducted excluding looks to the light effect for the Effect pair

10 (Eq. 2 for both pairs) which revealed a similar pattern: there were no differences between

1 conditions during the first time bin ( $p = .36$ ) but a significant difference emerged during the  
 2 third time bin ( $p = .001$ ).

### 3 3.2 Predicting target objects: correct vs. incorrect

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

$$11 \quad \text{Incorrect}_{\text{Object+light}} = \frac{\# \text{ looks to other 4 objects}/4}{\text{total \# looks to all objects+light}} \quad (3)$$

$$12 \quad \text{Incorrect}_{\text{Object}} = \frac{\# \text{ looks to other 4 objects}/4}{\text{total \# looks to all objects}} \quad (4)$$

13 An ANOVA with Prediction (*Correct, Incorrect*) and Pair (*Effect, No-effect*) as within-  
 14 subjects factors and Condition (*Agent, Ghost*) as a between-subjects factor revealed significant  
 15 main effects of Pair,  $F(1,54) = 8.08$ ,  $p = .006$ ,  $\eta_p^2 = .13$ , and Condition,  $F(1,54) = 5.95$ ,  $p = .018$ ,  $\eta_p^2$   
 16  $= .09$ , and a significant Condition by Prediction interaction,  $F(1,54) = 98.50$ ,  $p = .018$ ,  $\eta_p^2 = .10$ .  
 17 Planned comparisons to follow up on this interaction effect (Bonferroni-corrected critical  $p$ -  
 18 value calculated as .025) revealed that *Correct* proportions were significantly greater than  
 19 *Incorrect* proportions for toddlers in the *Agent* conditions ( $p = .003$ ) whereas there was no  
 20 difference for toddlers in the *Ghost* condition ( $p = .63$ ; see Table 1; Figure 4, left).

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

1 (Eq. 2&4 for both pairs). The same patterns emerged, with a significant interaction between  
 2 Condition (*Agent, Ghost*) and Prediction ( $Correct_{Object}, Incorrect_{Object}$ ),  $F(1,54) = 9.00$ ,  $p$   
 3  $= .004$ ,  $\eta_p^2 = .14$ .

4 Table 1

5 *Main dependent measures separated by condition*

DV	Pair	Factor	Agent		Ghost	
			Mean	SD	Mean	SD
Correct vs. Incorrect	Effect <sub>Object + Light</sub>	Correct <sub>Object + Light</sub>	0.31*	0.19	0.24*	0.14
		Incorrect <sub>Object + Light</sub>	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 <sub>Object</sub>	Correct <sub>Object</sub>	0.20*	0.16	0.11*
		Incorrect <sub>Object</sub>	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

6 *Note.* \* $p < .025$  (difference between Agent and Ghost conditions). Fewer participants contributed action  
 7 performance data, as some toddlers became too fussy to complete the play session following action observation.  
 8

### 9 3.3 Learning statistical context: deterministic vs. random transitions

10 Targets occurred both within deterministic pairs and also at random outside of these  
 11 pairs (see *Methods*, 2.2.1). For example, *open* occurred within a deterministic transition and  
 12 was 100% predictable after *bend* had occurred. *Open* also occurred after other events, but  
 13 featured low predictability (16.7%) at the onset of these (e.g., *slide*, *push*, *open*, or *spin*). We  
 14 defined random transitions as the events that preceded the targets when they occurred  
 15 randomly (Eq. 6) outside of their corresponding deterministic pair (i.e., a measure of ‘chance’,

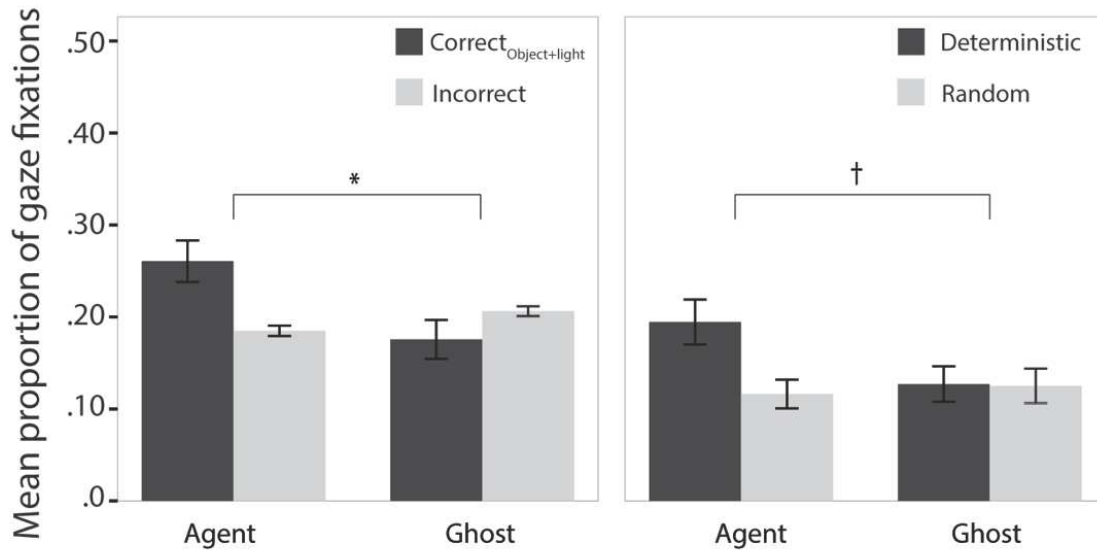
1 as defined in Tummeltshammer & Kirkham, 2013). We discarded trials in which the same action  
 2 repeated itself (for example: *push* followed by *push*) as it was impossible to determine whether  
 3 fixations during these trials were anticipations or the toddlers simply not moving their eyes.  
 4 Analyzing the difference between deterministic (Eq. 5; same as *Correct*, Eq. 2) and random (Eq.  
 5 6) transitions enabled us to compare fixations to the same target objects in different statistical  
 6 contexts. In these analyses, looks to the light effect were excluded entirely.

$$7 \quad \textit{Deterministic} = \frac{\# \textit{fixations to target (predictive trials)}}{\textit{total \# fixations to objects}} \quad (5)$$

$$8 \quad \textit{Random} = \frac{\# \textit{fixations to target (non-predictive trials)}}{\textit{total \# fixations to objects}} \quad (6)$$

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

17 Figure 4: Proportions of predictive fixations during action observation. *Left*: correct vs.  
 18 incorrect (Eqs. 1-4, including the light in proportion scores for the *Effect* pair). *Right*:  
 19 deterministic vs. random transitions (Eqs. 5-6). Bars represent standard errors. \* $p < .025$ , † $p$   
 20 = .06



1

## 2 3.4 Behavior Coding

3 Behavior during the familiarization phase and during the first two minutes of the play  
 4 session following the video session was coded for any instance in which the child made physical  
 5 contact with one of the objects on the stimulus. A *success* was noted whenever the toddler  
 6 made contact with an object and completed the action before removing his or her hand or  
 7 visibly pausing (e.g., opening and closing the yellow door). An *attempt* was noted any time the  
 8 child made contact with an object and completed part of the action before removing his or her  
 9 hand or made an overt effort to manipulate an object in some way without success (e.g.,  
 10 opening the yellow door, but leaving it open). If the child only placed his or her hand on an  
 11 object without attending to it or attempting to manipulate it (i.e. no attempt to perform an  
 12 ‘action’), these instances were considered invalid and were excluded from analyses. We  
 13 counted the number of Effect and No-effect pairs that each toddler reproduced and the total  
 14 number of actions they produced during the familiarization and post-observation phases. Fifty

1 percent of the total participants were coded by a second coder. There was strong agreement  
2 between the two coders, Cohen's  $\kappa = .729$ ,  $p < .001$ .

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

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

19       To assess the possibility that toddlers who were better at detecting the sequence  
20 regularities were simply more active to begin with, we assessed toddlers' baseline activity  
21 during the familiarization phase prior to action observation. Baseline activity (defined as the  
22 mean number of actions performed per minute) did not correlate toddlers' proportions of

1 correct fixations for either the Effect pair (including or excluding the effect) or the No-effect  
2 pair ( $ps > .62$ ).

### 3 4.0 Discussion

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

10 We evaluated three related measures of statistical learning that targeted different  
11 underlying questions regarding toddlers' improvement over trials and their overall predictive  
12 looking performance. All three measures demonstrated a similar pattern of results, with  
13 toddlers in the Agent condition demonstrating learning of the statistical structure in the  
14 sequence whereas those in the Ghost condition did not. Specifically, we found that only  
15 toddlers in the Agent condition showed a linear increase in their correct predictions for  
16 deterministic pairs over the course of the experiment. These toddlers also anticipated correct  
17 locations more frequently than incorrect locations during deterministic pairs, and made a  
18 greater rate of anticipatory fixations to target objects during deterministic trials than random  
19 trials. Toddlers in the Ghost condition, in contrast, did not differentiate between correct and  
20 incorrect locations and did not look to target objects more during deterministic transitions than  
21 during random transitions. Finally, for toddlers in the Agent condition, rates of correct  
22 predictions also correlated with their own spontaneous action performance following action



1 observation. Our findings thus provide new evidence that on-line statistical learning directly  
2 contributes to predictive gaze behaviors during toddlers' observation of continuous action  
3 sequences.

#### 4 4.1 Learning from observing an agent's actions vs. visual events

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

18       Social cues help infants notice statistical properties of objects in noisy visual scenes (Wu,  
19 Gopnik, Richardson, & Kirkham, 2010). With limited attention and cognitive resources, social  
20 cues likely serve as an adaptive filter to help toddlers and infants select relevant information  
21 from the environment. Although general attention to the displays did not differ between

1 conditions, it is possible that the presence of a social agent, as depicted by the hand, selectively  
2 directed anticipatory attention toward the sequential patterns in the ongoing actions.

3         The reason why cues such as hands and faces selectively guide infants' attention is an  
4 open question. In our paradigm, there were no pedagogical, facial, or other bodily cues present;  
5 observation of the actor's hand alone drove the effect observed between conditions. One  
6 possible explanation for this effect echoes current embodied accounts of the unique role of the  
7 motor system during action observation (Hunnius & Bekkering, 2014; Woodward & Gerson,  
8 2014). Infants become better at predicting actions as they acquire motor experience with the  
9 observed movements, suggesting that motor activity during observation facilitates their  
10 predictions (Kanakogi & Itakura, 2011; Stapel et al., 2016). Compelling evidence from adult  
11 research has shown that motor activation is causally related to action prediction (Elsner et al.,  
12 2013) and similar evidence has been found in infancy (Marshall, Young, & Meltzoff, 2011;  
13 Stapel et al., 2016; van Elk, van Schie, Hunnius, Vesper, & Bekkering, 2008). Our findings are  
14 consistent with the idea that the motor system facilitates action predictions during observation.  
15 Though we did not directly test motor activity here, it is a theoretically grounded possibility that  
16 we are currently pursuing in follow-up studies.

17         An alternative explanation for the different patterns found in the Agent versus Ghost  
18 conditions relies on low-level perceptual accounts. For example, the toddlers need not  
19 represent anything about actions or agents if the saliency of the events differed between  
20 conditions as a function of the use of a hand versus a spotlight. Although we accounted for  
21 looking time and number of fixations, it is possible that some unidentifiable difference in

1 saliency existed between conditions. If this were the case, however, we would not expect a  
2 relation between action prediction and reproduction specifically for the Agent condition.

### 3 4.2 Learning in infancy: structure-detecting skills

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

14         Despite providing evidence for learning, we found that toddlers' proportion scores were  
15 generally low and did not exceed 30% except in the Agent condition on later trials. One reason  
16 for this is likely due to the fact that, unlike previous studies, we only examined toddlers'  
17 anticipatory (rather than reactive) looks to upcoming targets. In a recent experiment  
18 investigating spatiotemporal learning in 8-month-olds, approximately 35% of infants' gaze  
19 fixations were anticipatory, whereas the remaining 65% were reactive (Tummeltshammer &  
20 Kirkham, 2013). For more complex stimuli (i.e., multiple objects presented simultaneously to  
21 which infants must distribute their visual attention), up to 95% of infants' fixations were  
22 reactive rather than predictive (Tummeltshammer, Pomiechowska, Gliga & Kirkham, 2015). In

1 the current study, we introduced the additional demands of tracking statistical relations  
2 between actions and events without spatial information to help guide predictions. We found  
3 that toddlers can make anticipatory eye movements for these complex sequences when they  
4 observed the Agent condition, but not the Ghost condition.

#### 5 4.3 The influence of action-effects on learning

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

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

#### 1 4.4 Action performance

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

18       For the toddlers who did not reproduce any action pairs, exploration may have simply  
19 been more interesting than activating the light or imitating the observed sequence. Exploration  
20 is a critical developmental behavior that allows discovery of new environmental features  
21 (Hopper et al., 2010). Though important in certain contexts, adult instruction causes children to  
22 limit their exploratory play in order to focus on performing target actions (Bonawitz, Shafto,

1 Gweon, Goodman, Spelke, & Schulz, 2011). The toddlers who did reproduce effect pairs may  
2 have interpreted the agent's demonstration as an instructional cue, and constrained their  
3 behavior to those actions more likely to cause a salient outcome (e.g., Bushnell & Boudreau,  
4 1998).

#### 5 4.5 Conclusion

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

18

19

## References

- 1  
2 Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: Active inference  
3 in the motor system. *Brain Structure and Function*, *218*(3), 611-643. doi: 10.1007/s00429-012-  
4 0475-5
- 5 Amso, D., & Davidow, J. (2012). The development of implicit learning from infancy to  
6 adulthood: item frequencies, relations, and cognitive flexibility. *Developmental*  
7 *Psychobiology*, *54*(6), 664-673. doi: 10.1002/dev.20587
- 8 Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic  
9 action. *Child Development*, *72*(3), 708-717. doi: 10.1111/1467-8624.00310
- 10 Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The  
11 double-edged sword of pedagogy: Instruction limits spontaneous exploration and  
12 discovery. *Cognition*, *120*(3), 322–330. doi: 10.1016/j.cognition.2010.10.001
- 13 Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children’s imitation of  
14 causal action sequences is influenced by statistical and pedagogical evidence. *Cognition*, *120*(3),  
15 331-340. doi: 10.1016/j.cognition.2010.12.001
- 16 Bushnell, E. W., & Boudreau, J. P. (1998). Exploring and exploiting objects with the hands  
17 during infancy. In Connolly, K. J. (Ed.), *The Psychobiology of the Hand. Clinics in Developmental*  
18 *Medicine*, *147*, 144-161. London, England: Mac Keith Press.
- 19 Canfield, R. L., & Haith, M. M. (1991). Young infants' visual expectations for symmetric  
20 and asymmetric stimulus sequences. *Developmental Psychology*, *27*(2), 198. doi: 10.1037/0012-  
21 1649.27.2.198

- 1 Danish, D., & Russell, J. (2007). The role of 'action-effects' and agency in toddlers'  
2 imitation. *Cognitive Development*, 22(1), 69-76. doi: 10.1016/j.cogdev.2006.08.001
- 3 Elsner, B., Hommel, B., Mentschel, C., Drzezga, A., Prinz, W., Conrad, B., & Siebner, H.  
4 (2002). Linking actions and their perceivable consequences in the human brain. *NeuroImage*,  
5 17(1), 364-372. doi: 10.1006/nimg.2002.1162
- 6 Elsner, C., D'Ausilio, A., Gredebäck, G., Falck-Ytter, T., & Fadiga, L. (2013). The motor  
7 cortex is causally related to predictive eye movements during action observation.  
8 *Neuropsychologia*, 51(3), 488-492. doi: 10.1016/j.neuropsychologia.2012.12.007
- 9 Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's  
10 action goals. *Nature Neuroscience*, 9(7), 878-879. doi: 10.1038/nn1729
- 11 Fawcett, T. W., Skinner, A. M. J., & Goldsmith A. R. (2002). A test of imitative learning in  
12 starling using two-action method with an enhanced ghost control. *Animal Behaviour*, 64, 547-  
13 556. doi: 10.1006/anbe.2002.3092
- 14 Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*,  
15 424(6950), 769-771. doi: 10.1038/nature01861
- 16 Gerson, S. A., & Woodward, A. L. (2013). The goal trumps the means: Highlighting goals is  
17 more beneficial than highlighting means in means-end training. *Infancy*, 18(2), 289-302. doi:  
18 10.1111/j.1532-7078.2012.00112.x
- 19 Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards:  
20 Dissociable neural prediction error signals underlying model-based and model-free  
21 reinforcement learning. *Neuron*, 66(4), 585-595. doi: 10.1016/j.neuron.2010.04.016



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

4 Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms  
5 in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of  
6 variation and covariation. *Developmental Psychology*, 37(5), 620-629. doi: 10.1037/0012-  
7 1649.37.5.620

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

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

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

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

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

- 1 Jovanovic, B., Kiraly, I., Elsner, B., Gergely, G., Prinz, W., & Aschersleben, G. (2007). The  
2 role of effects for infants' perception of action goals. *Psychologia* 50(4), 273-290. doi:  
3 10.2117/psysoc.2007.273
- 4 Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action  
5 prediction and motor ability in early infancy. *Nature Communications*, 2, 341. doi:  
6 10.1038/ncomms1342
- 7 Kenny, David A. (1987). *Statistics for the social and behavioral sciences*. Little, Brown.
- 8 Keyzers, C., Kaas, J. H., & Gazzola, V. (2010). Somatosensation in social perception. *Nature*  
9 *Reviews Neuroscience*, 11(6), 417-428. doi: 10.1038/nrn2919
- 10 Kirkham, N., Slemmer, J., & Johnson, S. (2002). Visual statistical learning in infancy:  
11 Evidence for a domain general learning mechanism. *Cognition*, 83(2), B35-B42.  
12 doi:10.1016/S0010-0277(02)00004-5
- 13 Klein, A. M., Hauf, P., & Aschersleben, G. (2006). The role of action effects in 12-month-  
14 olds' action control: A comparison of televised model and live model. *Infant Behavior and*  
15 *Development*, 29(4), 535-544. doi: 10.1016/j.infbeh.2006.07.001
- 16 Kochukhova, O., & Gredebäck, G. (2010). Preverbal infants anticipate that food will be  
17 brought to the mouth: An eye tracking study of manual feeding and flying spoons. *Child*  
18 *Development*, 81(6), 1729-1738. doi: 10.1111/j.1467-8624.2010.01506.x
- 19 Krogh, L., Vlach, H., & Johnson, S. P. (2013). Statistical learning across development:  
20 Flexible yet constrained. *Frontiers in Psychology*, 3, 598. doi: 10.3389/fpsyg.2012.00598

- 1 Marshall, P. J., Young, T., & Meltzoff, A. N. (2011). Neural correlates of action observation  
2 and execution in 14-month-old infants: An event-related EEG desynchronization study.  
3 *Developmental Science*, 14(3), 474-480. doi: 10.1111/j.1467-7687.2010.00991.x
- 4 Paulus, M., Hunnius, S., van Wijngaarden, C., Vrans, S., van Rooij, I., & Bekkering, H.  
5 (2011). The role of frequency information and teleological reasoning in infants' and adults'  
6 action prediction. *Developmental Psychology*, 47(4), 976-983. doi: 10.1037/a002
- 7 Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive*  
8 *Psychology*, 9(2), 129-154. doi:10.1080/713752551
- 9 Ray, E., & Heyes, C. (2011). Imitation in infancy: the wealth of the stimulus.  
10 *Developmental Science*, 14(1), 92-105. doi: 10.1111/j.1467-7687.2010.00961.
- 11 Reid, V. M., Hoehl, S., Grigutsch, M., Groendahl, A., Parise, E., & Striano, T. (2009). The  
12 neural correlates of infant and adult goal prediction: Evidence for semantic processing systems.  
13 *Developmental Psychology*, 45(3), 620-629. doi: 10.1037/a0015209
- 14 Romberg, A. R., & Saffran, J. R. (2012). Expectancy learning from probabilistic input by  
15 infants. *Frontiers in Psychology*, 3, 610. doi: 10.3389/fpsyg.2012.00610
- 16 Ruffman, T., Taumoepeau, M., & Perkins, C. (2012). Statistical learning as a basis for social  
17 understanding in children. *British Journal of Developmental Psychology*, 30(1), 87-104. doi:  
18 10.1111/j.2044-835X.2011.02045.x
- 19 Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants.  
20 *Science*, 274, 1926 - 1928. doi: 10.1126/science.274.5294.1926

- 1 Saylor, M. M., Baldwin, D. A., Baird, J. A., & LaBounty, J. (2007). Infants' on-line  
2 segmentation of dynamic human action. *Journal of Cognition and Development, 8*(1), 113-128.  
3 doi: 10.1080/15248370709336996
- 4 Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a  
5 new framework. *Trends in Cognitive Sciences, 11*(5), 211-218.
- 6 Stahl, A. E., Romberg, A. R., Roseberry, S., Golinkoff, R. M., & Hirsh-Pasek, K. (2014).  
7 Infants segment continuous events using transitional probabilities. *Child Development 85*(5),  
8 1821-1826. doi: 10.1111/cdev.12247
- 9 Stapel, J. C., Hunnius, S., van Elk, M., & Bekkering, H. (2010). Motor activation during  
10 observation of unusual versus ordinary actions in infancy. *Social Neuroscience, 5*(5-6), 451-460.  
11 doi: 10.1080/17470919.2010.490667
- 12 Stapel, J. C., Hunnius, S., Meyer, M., & Bekkering, H. (2016). Motor system contribution to  
13 action prediction: Temporal accuracy depends on motor experience. *Cognition, 148*, 71-78. doi:  
14 10.1016/j.cognition.2015.12.007
- 15 Thompson, D. E., & Russell, J. (2004). The ghost condition: Imitation versus emulation in  
16 young children's observational learning. *Developmental Psychology, 40*(5), 882. doi:  
17 10.1037/0012-1649.40.5.882
- 18 Tummeltshammer, K. S., & Kirkham, N. Z. (2013). Learning to look: Probabilistic variation  
19 and noise guide infants' eye movements. *Developmental Science, 16*(5), 760-771. doi:  
20 10.1111/desc.12064

- 1 Tummeltshammer, K., Pomiechowska, B., Gliga, T. & Kirkham, N. (2015). *Spatiotemporal*  
2 *regularity and infants' visual predictions: an ERP study*. Poster presented at the International  
3 Conference on Interdisciplinary Advances in Statistical Learning, San Sebastian, Spain.
- 4 Van Elk, M., van Schie, H. T., Hunnius, S., Vesper, C., & Bekkering, H. (2008). You'll never  
5 crawl alone: Neurophysiological evidence for experience-dependent motor resonance in  
6 infancy. *NeuroImage*, 43(4), 808-814. doi: 10.1016/j.neuroimage.2008.07.057
- 7 Waismeyer, A., Meltzoff, A. N., & Gopnik, A. (2015). Causal learning from probabilistic  
8 events in 24-month-olds: An action measure. *Developmental Science*, 18(1), 175-182. doi:  
9 10.1111/desc.12208
- 10 Wentworth, N., Haith, M., & Hood, R. (2002). Spatiotemporal regularity and inter-event  
11 contingencies as information for infants' visual expectations. *Infancy*, 3(3), 303-321. doi:  
12 10.1207/S15327078IN0303\_2
- 13 Woodward, A. L., & Gerson, S. A. (2014). Mirroring and the development of action  
14 understanding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1644),  
15 20130181. doi: 10.1098/rstb.2013.0181.
- 16 Woodward, A. L., & Sommerville, J. A. (2000). Twelve-month-old infants interpret action  
17 in context. *Psychological Science*, 11(1), 73-77. doi: 10.1111/1467-9280.00218
- 18 Wu, Rachel and Gopnik, A. and Richardson, D.C. and Kirkham, Natasha Z. (2010) Social  
19 cues support learning about objects from statistics in infancy. In Ohlsson, S. and Catrambone, R.  
20 (Eds.) *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (1228-1233).  
21 Austin, Texas, USA: Cognitive Science Society.

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