



UNIVERSITY  
OF WOLLONGONG  
AUSTRALIA

University of Wollongong  
Research Online

---

Faculty of Social Sciences - Papers

Faculty of Social Sciences

---

2018

# Changes in infant visual attention when observing repeated actions

Felix Koch

*Linköping University*

Anett Sundqvist

*Linköping University*

Jane S. Herbert

*University of Wollongong, herbertj@uow.edu.au*

Tomas Tjus

*University of Gothenburg*

Mikael Heimann

*Linköping University, Swedish Institute for Disability Research*

---

## Publication Details

Koch, F., Sundqvist, A., Herbert, J., Tjus, T. & Heimann, M. (2018). Changes in infant visual attention when observing repeated actions. *Infant Behavior and Development*, 50 189-197.

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: [research-pubs@uow.edu.au](mailto:research-pubs@uow.edu.au)

---

# Changes in infant visual attention when observing repeated actions

## **Abstract**

2018 Elsevier Inc. Infants' early visual preferences for faces, and their observational learning abilities, are well-established in the literature. The current study examines how infants' attention changes as they become increasingly familiar with a person and the actions that person is demonstrating. The looking patterns of 12- (n = 61) and 16-month-old infants (n = 29) were tracked while they watched videos of an adult presenting novel actions with four different objects three times. A face-to-action ratio in visual attention was calculated for each repetition and summarized as a mean across all videos. The face-to-action ratio increased with each action repetition, indicating that there was an increase in attention to the face relative to the action each additional time the action was demonstrated. Infant's prior familiarity with the object used was related to face-to-action ratio in 12-month-olds and initial looking behavior was related to face-to-action ratio in the whole sample. Prior familiarity with the presenter, and infant gender and age, were not related to face-to-action ratio. This study has theoretical implications for face preference and action observations in dynamic contexts.

## **Disciplines**

Education | Social and Behavioral Sciences

## **Publication Details**

Koch, F., Sundqvist, A., Herbert, J., Tjus, T. & Heimann, M. (2018). Changes in infant visual attention when observing repeated actions. *Infant Behavior and Development*, 50 189-197.

**Changes in infant visual attention when observing repeated actions**

Felix-Sebastian Koch and Anett Sundqvist

Linköping University, Sweden

Jane Herbert

University of Wollongong, Australia

Tomas Tjus

University of Gothenburg, Sweden

Mikael Heimann

Linköping University, Sweden

Author Note

Felix-Sebastian Koch, Infant and Child Lab, Department of Behavioural Sciences and Learning, Linköping University, Sweden; Anett Sundqvist, Infant and Child Lab, Department of Behavioural Sciences and Learning, Linköping University, Sweden; Jane Herbert, School of Psychology, University of Wollongong, Australia; Tomas Tjus, Department of Psychology, University of Gothenburg, Sweden; Mikael Heimann, Infant and Child Lab, Department of Behavioural Sciences and Learning, Linköping University, Sweden.

This research was supported by a grant from the Swedish Research Council (grant # 2011-1913).

Correspondence concerning this article should be addressed to Felix-Sebastian Koch, Infant and Child Lab Linköping, Department of Behavioural Sciences and Learning, Linköping University, SE-581 83 Linköping, Sweden. E-mail: felix.koch@liu.se

### **Abstract**

Infants' early visual preferences for faces, and their observational learning abilities, are well-established in the literature. The current study examines how infants' attention changes as they become increasingly familiar with a person and the actions that person is demonstrating. The looking patterns of 12- (n = 61) and 16-month-old infants (n = 29) were tracked while they watched videos of an adult presenting novel actions with four different objects three times. A face-to-action ratio in visual attention was calculated for each repetition and summarized as a mean across all videos. The face-to-action ratio increased with each action repetition, indicating that there was an increase in attention to the face relative to the action each additional time the action was demonstrated. Infant's prior familiarity with the object used was related to face-to-action ratio at in 12-month-olds and initial looking behavior was related to face-to-action ratio in the whole sample. Prior familiarity with the presenter, and infant gender and age, were not related to face-to-action ratio. This study has theoretical implications for face preference and action observations in dynamic contexts.

*Keywords:* visual attention; face preference; action observation; eye tracking

## 1. Introduction

One of the primary learning mechanisms for infants is observing what others are doing (Bandura, 1971; Meltzoff, Kuhl, Movellan, & Sejnowski, 2009). Naturalistic studies have shown that between the ages of 12- to 18-months, infants learn 1 to 2 new behaviors a day simply through observing the people around them (Barr & Hayne, 2003). In these complex learning situations, multiple sources of social and behavioral information are available to help the infant interpret and benefit from the events they observe, especially if they see the same event demonstrated multiple times. We examine here the factors that influence how infants distribute their attention to elements of a dynamic learning situation (a person's face and the actions that the person is producing) across time. Identifying how attention changes as events are repeated, and are becoming increasingly familiar, will provide a better understanding of the learning mechanisms that guide infant cognitive development.

Visual preference procedures which use static images have consistently found that infants attend longer to faces compared to other stimuli (e.g., pictures of faces compared to pictures of toys). Several studies have demonstrated that this effect is in place from birth for face-like stimuli (Fantz, 1963; Johnson, Dziurawiec, Ellis, & Morton, 1991). From 4- to 5-months of age, infants attend for longer to pictures of faces than distractor stimuli (Di Giorgio, Turati, Altoè, & Simion, 2012; Gliga, Elsabbagh, Andravizou, & Johnson, 2009; Gluckman & Johnson, 2013; Libertus & Needham, 2011; DeNicola, Holt, Lambert & Cashon, 2013), and attentional bias towards faces becomes a robust effect thereafter (e.g. Amso, Haas, & Markant, 2014; Kwon, Setoodehnia, Baek, Luck, & Oakes, 2014; Leppänen, 2016). Following from Cohen (1972) this is often referred to as the attention holding effect of faces (e.g. DeNicola et al.).

When presented with a dynamic context, attention to faces has also been shown to increase during the first year of life (Frank, Amso, & Johnson, 2014; Frank, Vul, & Johnson, 2009) and then remains present throughout life (Stoesz & Jakobson, 2014). Furthermore, Frank et al. (2014) reported that infants' attentional abilities in general are related to how much they attend to faces. Infants who were quicker and more accurate to identify targets in visual search tasks also looked longer at faces when viewing dynamic scenes. Although faces are of prime interest to infants aged 3 to 30 months when viewing dynamic stimuli, with age infants increase their attention to what a person is doing, with older infants attending relatively more to the person's hands than younger infants do (Frank, Vul, & Saxe, 2012).

Changes in volitional control of attention (Colombo, 2001; Courage & Setliff, 2010) may play a role in age-related changes in infant attention to, and memory for, aspects of dynamic scenes. Bahrick and Newell (2008) presented infants with videos of adults demonstrating everyday activities (hair brushing, teeth brushing, blowing bubbles, or applying make-up) and tested memory for the faces and actions using a novelty preference test. While 5.5-month-old infants showed memory for the action being performed, by 7 months of age infants showed memory for both the performers' faces and their actions. The authors argued that actions are more salient than the presenters' faces and 5-month-olds do not have the attentional resources to register both the action and the face. From the age of 7 months, infants have the resources to register both elements (Bahrick, Gogate, & Ruiz, 2002). Using eye-tracker methodology, Taylor and Herbert (2013, 2014) found that infants from 6 to 12 months of age attend less to the background and focused both on the presenter and the action she was performing, but did not find differences in attention for the presenter and the action. There is also evidence that 12-month-old infants (Kolling, Óturai, & Knopf, 2014) and 18-month-old infants (Óturai, Kolling,

& Knopf, 2013) attend more to actions than to the presenter's face, independent of whether the action with the object was functional or arbitrary. These eye-tracker studies analyzed infants' attention to the video during different periods but did not compare between repetitions of actions. Thus they did not consider how attention might change across the learning situation. Changes in infants' focus of attention over time when they are viewing novel actions could be important for understanding observational learning processes.

The current study aims to identify the relative distribution of infants' attention to a presenter's face and the repeatedly demonstrated actions. We consider two alternative predictions for how infants might distribute their visual attention over time. One alternative, in line with infants' primary interest in faces, is that infants will first attend to the presenter's face until they have sufficiently processed the social information, before then directing their attention to what that person is doing. This prediction would suggest that the relative distribution of visual attention to the face would decline over time, and visual attention to the action area would increase over time. An alternative suggestion comes from the research reviewed above on infants' action observation, and predicts that infants would first be interested in the action itself and only later shift their attention to the person performing the action. According to this account the relative distribution of visual attention to the face would increase as actions are repeated.

The reviewed literature shows infants' primary interest in faces, on the one hand, and their strong interest for action observation, on the other hand. None of the reviewed studies were designed to answer a direct question of whether infants prefer to look at faces or at actions. Such a comparison would depend very much on the context, and in particular the social context, of the action presentation. We focus here on the dynamics of where infants distribute their attention during the demonstration phase of the imitation paradigm, while the infant is observing an adult

demonstrate an action or sequence of actions with a novel object. Infants' imitation performance increases as a function of age (for review see Hayne, 2004), and additional demonstrations of target actions improves learning from a 2D televised presentation at all ages tested between 12- and 21-months (Barr, Muentener, Garcia, Chavez, & Fujimoto, 2007). Attentional mechanisms that may lie behind the effect have not been investigated. By comparing across two ages (12- and 16-months) we examine whether age might influence the observed distribution of visual attention across the repetitions, in line with increasing endogenous control of attention (Colombo, 2001).

The decline or increase in attention to the face relative to attention to the action may also be influenced by early gender differences in attending social stimuli. With 6-month old infants, Gluckman and Johnson (2013) have shown that social stimuli (faces, body parts, and animals) attract attention in a stimulus array compared to common objects. However, for girls especially, faces were the strongest attention holder. Furthermore, Mundy et al. (2007) reported that girls, slightly more than boys, used gaze and gestures to elicit aid from a social partner in a live interaction. Although the use of a pre-recorded video presentation would reduce the strength of social cues, the research mentioned above suggests that girls might attend more than boys to a presenter's face rather than her actions, either throughout or at some parts of the presentation. A third factor that may influence the distribution of visual attention is familiarity with the person or the object involved. Well-established findings of visual preference (for review see Rose, Feldman, & Jankowski, 2004) suggest that familiarity of a person or an object may influence how infants distribute their attention towards that person or object. Familiarity is usually shown by more attention to novel objects compared to familiar ones. Familiarity, in the current study, is established for some infants in real life before they are shown pre-recorded videos that show the

presenter or object they have been familiarized with. Due to the visual preference effect, infants might spend more time visually exploring the novel aspects (novel face or novel object) during observational learning, which would be indicated by a lower face-to-action ratio for infants who are familiar with the presenter compared to infants who are unfamiliar with the presenter, and a higher face-to-action ratio for infants who are familiar with the object compared to infants who are unfamiliar with the object. Finally, we examine the influence of differences in the microstructure of visual behavior. Jankowski, Rose, and Feldman (2001) studied 5-month-old infants' visual behavior with the visual paired comparison paradigm and found that infants with fewer shifts and longer looks at encoding did not show novelty preference whereas infants with more shifts and shorter looks during encoding did show novelty preference. These findings suggest that the microstructure of visual behavior could be related to visual processing and learning, as more frequent shifts and shorter looks relate to faster processing and learning as indicated by novelty preference. Gredebäck and Daum (2015) point out the importance of analyzing the temporal microstructure in visual behavior in dynamic settings in order to understand infants' processing of social stimuli. With the help of eye tracking technology, infants' microstructure in visual behavior can be analyzed. Using this method, Papageorgiou, Smith, Wu, Johnson, Kirkham, and Ronald (2014) found that mean fixation duration in the first year of life was related to parental reports of attentional and behavioral control for 3- to 4-year-old children. If infants vary in visual behavior when initially exploring the face or other parts of the stimuli they may show different patterns of visual attention during action observation.

In the current study, infants' relative distribution of visual attention between the face of the presenter and the action she performs were analyzed. Actions were demonstrated three times consecutively and dynamic changes in infants' attention were analyzed for each repetition. The

relative distribution of visual attention was analyzed by a face-to-action ratio that was calculated for each demonstration separately. It is hypothesized that the face-to-action ratio will change between demonstrations but no prediction is made for the direction of change. As discussed above it is plausible that infants first focus more on the person performing an action and later shift to observe more the action just as it is plausible that infants first attend more to the action and later more to the person doing the action.

The current study also explores whether gender, age, familiarity and microstructure in initial visual behavior relate to face-to-action ratio when observing repeated actions. Effects of gender and age were examined in a model including all infants tested at 12 and 16 months of age. Familiarity of the person and of the object was experimentally manipulated for 12-month-old infants. Microstructure in initial visual behavior was examined by classifying infants as using short and long fixations based on the peak fixation durations during the greeting phase before the actions were presented. It was hypothesized that there will be differences in face-to-action ratios for age, gender, familiarity, and microstructure in initial visual behavior as well as interactions between these factors and changes in face-to-action ratio between demonstrations.

## 2. Method

### 2.1. Participants

In the current cross-sectional study, 61 infants were 12 months of age ( $M = 368.9$  days,  $SD = 7.1$ ) and 29 infants were 16 months of age ( $M = 476.0$  days,  $SD = 6.7$ ). At 12 months of age, 32 infants were female (52.5 %) and all infants were born gestational week 35 or later ( $M$  g.a. = 40.2,  $SD = 1.5$ ), with a mean birth weight of 3708 g ( $SD = 542$ ) and birth length of 50.8 cm ( $SD = 2.3$ ). Most infants (91.8%) grew up in a monolingual Swedish-speaking household and had parents who had a university degree (77.1% of mothers and 52.5 % of fathers). At 16-

months, 20 infants were female (69 %) and all infants were born gestational week 37 or later ( $M$  g.a. = 40.0,  $SD = 1.3$ ), with a mean birth weight of 3770 g ( $SD = 446$ ) and birth length of 50.8 cm ( $SD = 1.9$ ). Most infants (82.1%) grew up in a monolingual Swedish-speaking household and had parents who had a university degree (86.2 % of mothers and 55.2 % of fathers).

Infants were included in the analyses if they provided fixation data (at least three seconds of a whole video clip and at least one fixation in the face or action area) for at least three of the four video clips analyzed in the current study. In total, attrition was 7 infants, all were 12 months old. One infant was tested but did not provide any data, two infants provided data for one video only and another four infants provided data for two videos only.

## **2.2. Procedure**

All infants were tested at the Infant and Child Lab at Linköping University, at a time of day that the parent reported as the infant's awake and alert period. Parental informed consent was obtained before testing. The parent and the infant met the experimenter a short walk away from the lab. The warm-up period for the infant to the experimenter was initiated during this walk, with smiles directed at the infant, although the experimenter primarily talked to the accompanying parent. Informed consent and background demographic information were obtained from the parent once they were in the lab, during which time the infant was free to explore the environment. The experimenter then began to interact more directly with the infant, smiling and handing him or her toys (none of the toys used during warm-up were used as stimuli in the study). When the infant showed signs of comfort, such as smiles or positive vocalizations, the experimental procedure was started. The infant was seated on their parent's lap in front of a Tobii T120 monitor (Stockholm, Sweden). A 36 seconds long infant friendly video clip (Baby Einstein) was used to attract attention to the screen. While the infant watched the video, the

distance from the monitor to the infant was adjusted to approximately 63 cm, with the monitor centered in front of the infant's face. When the video finished, the experimenter started the Tobii studio calibration procedure (five calibration points). After successful calibration, six experimental videos were shown. Parents were allowed to watch the videos together with their infant but were asked not to comment on what they saw or interfere with their infant's watching, other than to redirect the child to the screen if necessary. Parents were positioned in such a way that their eyes were outside of the virtual Tobii tracking box and the tracking status was monitored continuously online in order to detect any discrepancies from tracking the infant's eyes. No such discrepancies were observed.

The experimenter for 30 infants (attrition: 2 infants) at 12-months and 23 infants at 16-months was the same female person that also acted as the presenter in the video. An implication of this procedure was that some of the infants were familiar with the presenter in the video from real life. All other infants were tested by a male experimenter and were unfamiliar with the female presenter in the video. Furthermore, 13 infants (attrition: 3 infants) at age 12-months and two infants at age 16-months were allowed to play with each toy before they watched the video demonstrating the action on that object, in order to create familiarity with the object.

All infants had the opportunity to play with the object that was shown in the video after the video was finished. Infants played with each object for approximately one minute before their attention was attracted to the monitor again and the next video was shown. This was repeated until all six videos were shown. Recalibration occurred after three videos to prevent any drift in accuracy. Each infant saw the videos in one of six different orders. A Latin-square counterbalancing was used for the creation of the different orders. As described below, eye tracking data is analyzed for only four of the six videos.

### 2.3. Material

In the current study six 39-48 seconds long video clips were used, each portraying an adult presenter demonstrating single or multiple actions on an object. However, only four of the videos were relevant to the research question discussed here (as the action area in the remaining two videos could not be separated from the face area). Of the four videos included here, three showed single actions with an object and one showed multiple actions demonstrated with a hand-held puppet. All videos showed the same female presenter seated behind a beige wooden table facing the camera (see Figure 1). The background in the videos was a white wall without specific features. The presenter kept a happy animated tone throughout the video to keep the infants' attention directed to the screen. The videos followed the same general sequence, and all started with the presenter waving at the camera and using common Swedish greetings suitable for children and infants. Initially, the target objects for each task were visible on the left side of the screen. After a few seconds the presenter focused on the object to be used by saying "look at this" or similar phrases and placed the object in front of her, in the middle of the screen. Then she demonstrated the specific action that could be performed with that object (e.g., putting a string of beads in a cup; showing how a telescope extendable cup could be collapsed by pressing on it; shaking a blue toy egg that produces a rattle sound). For the multiple action task, the presenter held the puppet in her right hand and kept the puppet on the left side of the screen (see Figure 1) throughout the demonstration of the following actions: (1) removing a mitten from the right arm of the hand-held puppet, (2) shaking the mitten (causing a jingle bell attached inside to ring) and (3) putting the mitten back on the puppet's right arm. The target actions were demonstrated three times each for the single action tasks and the sequence of three target actions was presented three times for the multiple action task.

Before demonstrating or repeating each action, the presenter put both hands on the table (in the video with the hand-held puppet, only her right hand). The placement of the hands on the table was used for separating the different segments of the video: the greeting phase, first, second, and third repetition of the target actions. Mean duration of the video segments in seconds (with *SD* in parentheses) was 8.44 (2.13), 10.96 (1.45), 13.02 (2.20), 10.45 (2.80), for the greeting and the first, the second, and the third demonstration, respectively.

#### **2.4. Eye tracking data**

Eye tracking data was collected at 120 Hz with a Tobii T120 while infants watched the stimuli videos. The videos were presented through Tobii Studio (Tobii, Stockholm, Sweden), which was also used for calibration and data analyses. Figure 1 shows a screen shot of one of the stimuli videos (hand-held puppet) with borders of action and face areas of interest highlighted. Based on a distance of 63 cm from the screen, the size of the stimuli video in visual degrees was  $28.1^\circ \times 16.8^\circ$  angle. The action area was rectangular and extended  $14.9^\circ \times 12.7^\circ$  visual angle. The face area was oval and extended  $4.3^\circ \times 5.2^\circ$  visual angle. As can be seen in Figure 1, the red oval of the face area overlaps with the orange rectangle. Fixation time within the overlap counted only towards fixation time in the face area. The exact same location and dimensions for face and action area were used for analysis of all four stimuli videos. Eye tracking data was collected from the start to the end of each video. No attention-getting stimuli were used before showing each stimuli video.

<< Note: Insert Figure 1 about here >>

#### **2.5. Data reduction and statistical analysis**

The dependent variable in the current study is face-to-action ratio, which was calculated by dividing fixation time in the face area with the sum of fixation time in the face area and fixation time in the action area. The reported face-to-action ratio is a mean across four video clips, one for each video segment. Thus each infant's data is summarized in four ratio means.

In order to examine infants' microstructure in initial visual behavior, infants were classified as using longer or shorter peak fixations. During a fixation a spatial location is continuously in focus and a fixation ends with a saccade, a sudden change in spatial location. Fixations reported here were defined by the Tobii fixation filter included in Tobii Studio (Tobii, Stockholm, Sweden). The peak fixation is the longest single fixation in a spatial location. Each of the four videos analyzed had its own greeting phase before the action was demonstrated and a mean for peak fixation duration was calculated for each infant based on peak fixation duration measured in the greeting phase of each video. Peak fixation durations in the face area were significantly longer than in the action area (Table 1), which is in line with the attention holding effect of faces (Cohen, 1972, DeNicola, et al., 2013). However, peak fixation durations in the face area and in the action area were not significantly correlated (age 12-months:  $r = .14$ ,  $p = .33$ ,  $n = 54$ , age 16-months:  $r = -.03$ ,  $p = .86$ ,  $n = 29$ ). This indicates that longer peak fixations in the face area do not indicate longer peak fixations in the action area and that infants' visual behavior may not be constant across types of objects they fixate. Therefore, infants were classified once based on a median split for peak fixation duration in the face area and once based on a median split for peak fixation duration in the action area. The reason for two median splits was to examine if the microstructure in visual behavior is constant or differs across kind of stimuli (observing the face vs. observing objects). There is an aged difference for peak fixation in the action area,  $t(81) = 2.93$ ,  $p < .01$ , but not in the face area,  $t(81) = 1.00$ ,  $p = .32$ . In order to not

confound the analysis of peak fixation with age, each age group is divided by median split within its own age group. The median for peak fixation duration in the face area was 1.519 secs for 12-month-olds and 1.600 secs for 16-month-olds. The median for peak fixation duration in the action area was 0.881 secs for 12-month-olds and 1.108 sec for 16-month-olds.

<< Note: Insert Table 1 about here >>

Repeated measures ANOVAs were used to analyze the change over time in face-to-action ratio. IBM SPSS statistics version 23.0.0.2, 64-bit edition, was used to run all the statistical analyses reported. An  $\alpha \leq .05$  was used as a cut-off for statistical differences and effect size is reported as  $\eta_p^2$ . Due to problems with sphericity according to Mauchly's test in some of the reported models a Greenhouse-Geisser correction for degrees of freedom was used and the correction factor Greenhouse-Geisser  $\epsilon$  is reported. Residual plots for the models were inspected visually and no model fit problems were observed.

## **2.6. Ethics**

Approval of the present study was granted by the Regional Ethical Review Board, Linköping, Sweden. Families did not receive any compensation for participation.

## **3. Results**

The main analysis compares face-to-action ratio across the different segments of the video clips for changes over time. A repeated measures ANOVA was performed with the mean for face-to-action ratio as dependent variable and the four segments of the videos as the independent variable. Gender and age (12 vs 16 month) were included as between group factors. The two-way interactions between gender and video segments, age and video segment, and gender and

age were also included in the model, as was the three-way interaction between gender, age, and video segment. Face-to-action ratio differed significantly between video segments,  $F(2.5, 197.09) = 47.4, p < .001, \eta_p^2 = .38, \epsilon = .83$ . There was no significant effect of gender,  $F(1, 79) = 0.18, p = .67, \eta_p^2 = .002$ , or age,  $F(1, 79) = 1.89, p = .17, \eta_p^2 = .02$ , and no significant interaction,  $F(1, 79) = .87, p = .36, \eta_p^2 = .01$ . Furthermore, no significant interactions were found between gender and video segment,  $F(2.5, 197.1) = 2.05, p = .12, \eta_p^2 = .03, \epsilon = .83$ , or between age and video segment,  $F(2.5, 197.1) = 0.44, p = .69, \eta_p^2 = .006, \epsilon = .83$ . The three-way interaction between gender, age and video segment was also non-significant,  $F(2.5, 197.1) = 1.21, p = .30, \eta_p^2 = .02, \epsilon = .83$ . Mean values are presented in Figure 2. Tests of within-participant contrasts indicate differences between greeting and first demonstration,  $F(1, 79) = 85.5, p < .001, \eta_p^2 = .52$ , between first and second demonstration,  $F(1, 79) = 12.5, p = .001, \eta_p^2 = .14$ , and between second and third demonstration,  $F(1, 79) = 42.2, p < .001, \eta_p^2 = .35$ . The face-to-action ratio decreased from the greeting phase to the first demonstration, but then increased from the first to the second demonstration and again from the second to the third demonstration. These analyses show that face-to-action ratio is sensitive to repetition of actions and that the face-to-action ratio increases with number of repetition. Furthermore, these analyses did not indicate any differences between girls and boys, nor between 12- and 16-month-olds, and two-way and three-way interactions were not found to be significant. Therefore, age and gender are not included as factors in the following models.

<< Note: Insert Figure 2 about here >>

### 3.1. Factors influencing face-to-action ratio over time

### 3.1.1. Familiarity with the presenter or object.

Familiarity was tested systematically only at the 12-month observation. First a model was constructed to test the effect of familiarity with the presenter that included all infants tested at 12 months. For 28 infants the presenter was the experimenter (and therefore familiar from real life) and for 26 infants the presenter was unknown. A repeated measures ANOVA was performed with the mean for face-to-action ratio as dependent variable and the four video segments as the independent variable. Familiarity with the presenter was included as a between participant factor. The main effect for video segment remained,  $F(2.4, 126.8) = 27.37, p < .001, \eta_p^2 = .35, \varepsilon = .81$ , but familiarity with the presenter did not account for further variance,  $F(1, 52) = 1.34, p = .25, \eta_p^2 = .03$ , and neither did the interaction between presenter familiarity and video segments,  $F(2.4, 126.8) = 0.74, p = .50, \eta_p^2 = .01, \varepsilon = .81$ .

Next, a model was constructed that examined object familiarity at 12 months. For this model, the 10 infants that were allowed to play with the objects were compared to the 44 infants that were unfamiliar with the objects. The repeated measures ANOVA used the mean for face-to-action ratio as dependent variable and the four video segments as the independent variable. Familiarity with the objects was included as a between participant factor. The main effect for video segment remained unchanged,  $F(2.4, 126.0) = 15.8, p < .001, \eta_p^2 = .23, \varepsilon = .81$ . Examining object familiarity indicated a main effect,  $F(1, 52) = 4.16, p = .047, \eta_p^2 = .07$ , but no interaction effect between object familiarity and video segments,  $F(2.4, 126.0) = 0.34, p = .75, \eta_p^2 = .01, \varepsilon = .81$ . A higher face-to-action ratio was observed for infants who were familiar with the objects used in the videos compared to infants who were not. As the difference between 12 and 16 months of age was non-significant, the same model was run including all infants from both ages, comparing 12 infants that were familiar with the objects to 71 infants that were not. The main

effect for video segment remained unchanged,  $F(2.5, 200.9) = 22.8, p < .001, \eta_p^2 = .22, \varepsilon = .83$ . However, object familiarity indicated only a trend,  $F(1, 81) = 3.73, p = .057, \eta_p^2 = .04$ , and, as previously, no interaction effect between object familiarity and video segment,  $F(2.5, 200.9) = 0.44, p = .69, \eta_p^2 = .005, \varepsilon = .83$ , was found. Thus, collapsing the data across age indicates the effect to a weaker degree, than data from 12-month-old infants only.

### 3.1.2. Microstructure in initial visual behavior.

Differences between infants using shorter versus longer peak fixations were first tested based on peak fixation duration in the face area (Table 2). The repeated measures ANOVA used, as above, the mean for face-to-action ratio as dependent variable and the four video segments as the independent variable. Due to the difference in peak fixation duration between 12 and 16 months of age, a median split was used to create two groups at each age based on the infants' microstructure in the initial visual behavior when looking at the face. This repeated measures ANOVA included all infants, comparing infants with shorter peak fixation duration ( $n = 41$ ) with infants with longer peak fixation duration ( $n = 42$ ) as measured in the face area, and this was entered as a between participant factor.

Video segment was significant,  $F(2.53, 205.0) = 52.68, p < .001, \eta_p^2 = .39, \varepsilon = .84$ , and the microstructure in initial visual behavior was also significant,  $F(1, 81) = 10.1, p = .002, \eta_p^2 = .11$ . A significant interaction effect between visual behavior and video segment was observed,  $F(2.53, 205.0) = 4.1, p = .01, \eta_p^2 = .05, \varepsilon = .84$ . Tests of within-participant contrasts indicated significant interactions from the greeting segment to the first demonstration,  $F(1, 81) = 7.3, p < .01, \eta_p^2 = .08$ , and from the first to the second demonstration,  $F(1, 81) = 6.1, p = .02, \eta_p^2 = .07$ , but not from the second to the third demonstration,  $F(1, 81) = 0.04, p = .85, \eta_p^2 < .001$ . The

group with longer peak fixation duration during the greeting phase dropped more in face-to-action ratio from the greeting to the first demonstration, and increased less from the first to the second demonstration, nevertheless maintaining a higher face-to-action ratio for each video segment compared to the group with shorter peak fixation duration in the face area.

<< Note: Insert Table 2 about here >>

Peak fixation duration in the action area was significantly shorter but not correlated with peak fixation duration in the face area. Therefore, differences between infants that used shorter vs longer peak fixations were also tested based on peak fixation duration in the action area (Table 2). As previously, the median split was performed for each age group separately, but the repeated measures ANOVA included all infants from both age groups, with face-to-action ratio as the dependent variable, video segment as the independent variable. Microstructure in visual behavior based on the action area was entered as the between participant factor and the interaction between visual behavior and video segment was also included. Video segment was significant,  $F(2.48, 201.2) = 50.39, p < .001, \eta_p^2 = .38, \epsilon = .83$ , as was the between participant factor microstructure in initial visual behavior (based on the action area),  $F(1, 81) = 23.4, p < .001, \eta_p^2 = .22$ . No significant interaction effect between visual behavior and video segment was observed,  $F(2.48, 201.2) = 0.6, p = .58, \eta_p^2 < .01, \epsilon = .83$ . Infants with longer peak fixation duration in the action area maintained a lower face-to-action ratio throughout the videos, compared to infants with longer peak fixation duration as measured in the action area during the greeting phase. **4.**

## Discussion

Dynamic changes were observed in infants' attention to a presenter and their actions across time. In all segments of the video clips, infants paid considerable attention to the presenter's face, as might be expected from previous studies when overall looking time has been calculated (Frank et al., 2014, 2009, 2012; Stoesz & Jakobson, 2014). However, infants increased attention to the face relative to the action with each additional demonstration of the action. During the first demonstration, attention focused more on the action being presented, after which attention slowly shifted back to the presenter's face. Infants increased their relative distribution of visual attention to the face as actions were repeated. A primary interest in the action relative to the face is in line with studies on action observation (Bahrick & Newell, 2008; Kolling et al., 2014; Óturai et al., 2013). However, the current study further suggests that the primary interest in the actions relative to the face is a temporary phenomenon and decreases over time. The observed increase in face-to-action ratio across action demonstrations was independent of the other factors examined. Infants increased their attention to the face with each additional demonstration irrespective of whether they used longer or shorter peak fixation durations, were females or males, were unfamiliar or familiar with the presenter or object used, or were 12- or 16-months old.

Some unexpected results were found in relation to infants' microstructure in initial visual behavior. First, the classification yielded different results depending on whether the face or the action area was used for identification of the peak fixations. Peak fixations in the face area were significantly longer than in the action area, as may be expected according to the attention holding effect of faces (Cohen, 1972, DeNicola et al, 2013), but there was no correlation of peak fixations in the face area and the action area. This indicates the microstructure of infants' visual behavior is not consistent across stimuli and therefore this factor was analyzed twice, once based

on visual behavior exhibited when looking at faces and once when looking at the action area. Infants who showed longer initial peak fixation duration when looking at faces had a higher face-to-action ratio overall compared to infants that showed shorter initial peak fixation duration when looking at faces. The opposite was found for peak fixation when looking at the action area. Longer initial peak fixations here were related to a lower face-to-action ratio overall. Thus longer initial peak fixation duration does not lead in general to a higher face-to-action ratio, which could be expected as the attention holding effect may be stronger for infants that have longer initial peak fixations. Rather our results suggest that it is important to take into account what kind of stimulus (presenter's face or the action demonstration only) infants attend to when the microstructure of initial visual behavior is assessed. An interaction was found between the microstructure in visual behavior and face-to-action ratio when peak fixation duration was assessed in the face area where infants classified as using longer fixation durations looked longer at the face during the greeting phase and then showed a steeper drop from the greeting to the first demonstration and a lower increase from the first to the second demonstration in face-to-action ratio than did infants that used shorter fixation durations. As infants were classified according to the peak fixation duration in the face area during the greeting phase, it is not surprising that there was a steeper drop in face-to-action ratio from the greeting phase. However, the main finding with this analysis is that infants using longer fixations in the face area initially, showed a higher face-to-action ratio for each demonstration of the actions.

Object familiarity was experimentally manipulated by giving some infants the opportunity to play with the objects before watching the videos. Infants with prior experience of the object used had a higher face-to-action ratio in all video segments compared to infants without the experience. Evidence for this was found at 12 months of age. After collapsing the

data over the two age groups, this effect was weakened to a trend. This might be related to the study design, as data collection at 16 months was not designed to examine this factor. It seems, that object familiarity can affect distribution of visual attention in action observation, as some aspects of novelty preference impact the face-to-action ratio. Infants for whom the object is novel spent more time looking at that object (relative to the face) than infants for whom the object is familiar.

Analyses of presenter familiarity did not reveal any significant results. It seems that familiarity with the presenter did not diminish the attention holding effect of faces in the way the familiarity with objects may have diminished the novelty preference for these objects. Future research could consider how face-to-action attention patterns change if the presenter subsequently introduces a new action with the familiar object. We predict that an initial increase in attention to the action area would again be followed by increasing attention to the face region across demonstrations.

#### **4.1. Methodological discussion**

The quality of the eye tracking data is always an issue when this technology is used (Gredebäck, Johnson, & von Hofsten, 2010; Oakes, 2012). The current study lacks an independent check of the calibration procedure, which has been used in some studies (e.g. Frank et al, 2012), but is not yet common practice in infant eye tracking studies (Oakes, 2012). Once the calibration procedure provided sufficient data for four of the five calibration points the calibration was accepted. Whether the Tobii calibration procedure provided correct data for where infants looked was not checked. For this reason, the action area was defined generously including all non-face parts of the body and where the action was performed. This allows a constant size of the action area across videos. Also, this decreased the need for accuracy in

calibration which would lead to higher attrition. Furthermore, due to drift the coordinate estimation of the infants' fixations may have been more correct in the beginning, just after calibration, than towards the end (Wass, Forssman, & Leppänen, 2014). Our findings indicate that infants reoriented towards the presenter's face the longer they watched the video. As the face area was a small area of interest, poor calibration and decreased accuracy over time would lead to an underestimation of the effect found here. If the measurement error increased over time this would contribute to data indicating that infants are not looking at the face when in fact they are. Thus the main finding of the current study is not undermined by this possibility, but future studies could describe the effect with more accuracy.

Due to the lack of an independent calibration check, it cannot be determined whether different groups of infants potentially had better or worse calibrations and quality of data than other groups. The findings regarding microstructure of visual behavior could be questioned on these grounds as infants with shorter peak fixations were found to be different to infants with longer peak fixations. Calibration precision could relate to registration of different visual behavior but this cannot be tested in the current study. However, the main finding of this study is based on a within-participant effect, namely the increase in face-to-action ratio over repetitions of the same action, and not on a between-participant effect that could be confounded by variations in calibration precision between groups.

## **5. Conclusions**

The current findings suggest a dynamic change in the distribution of infants' attention to a presenter's face and the action she performs. Infants attend more to the action during the first demonstration but reorient towards the face on the following demonstrations. Future research should examine the mechanism behind the reorientation to the face, as this would yield further

insights into the attention holding effect of faces. As observational learning occurs in a social context the reason for infants reorienting to the face may be driven by social interest and attempts to understand the presenter's intentions. Therefore, the reorientation to the face during action observation may be important for infants' learning processes.

**Acknowledgements**

We are grateful to the families who participated in this study and would like to thank Angelica Edorsson for help with data collection.

### References

- Amso, D., Haas, S., & Markant, J. (2014). An eye tracking investigation of developmental change in bottom-up attention orienting to faces in cluttered natural scenes. *PLoS ONE*, *9*, 1–7.
- Bahrnick, L. E., Gogate, L. J., & Ruiz, I. (2002). Attention and memory for faces and actions in infancy: The salience of actions over faces in dynamic events. *Child Development*, *73*, 1629–1643.
- Bahrnick, L. E., & Newell, L. C. (2008). Infant discrimination of faces in naturalistic events: Actions are more salient than faces. *Developmental Psychology*, *44*, 983–996.
- Bandura, A. (1971). *Social learning theory*. New York: General Learning Press.
- Barr, R., & Hayne, H. (2003). It's not what you know, It's who you know: Older siblings facilitate imitation during infancy. *International Journal of Early Years Education*, *11*, 7–21.
- Barr, R., Muentener, P., Garcia, A., Chavez, V., & Fujimoto, M. (2007). The effect of repetition on imitation from television during infancy. *Developmental Psychobiology*, *49*, 196–207.
- Cohen, L. B. (1972). Attention-getting and attention-holding processes of infant visual preferences. *Child Development*, *43*, 869–879.
- Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*, 337–367.
- Courage, M. L., & Setliff, A. E. (2010). When babies watch television: Attention-getting, attention-holding, and the implications for learning from video material. *Developmental Review*, *30*, 220–238.
- Di Giorgio, E., Turati, C., Altoè, G., & Simion, F. (2012). Face detection in complex visual

- displays: An eye-tracking study with 3- and 6-month-old infants and adults. *Journal of Experimental Child Psychology*, *113*, 66–77.
- DeNicola, C.A., Holt, N.A., Lambert, A.J. and Cashon, C. H. (2013) Attention-orienting and attention-holding effects of faces on 4- to 8-month-old infants. *International Journal of Behavioral Development*, *37*, 143-147.
- Fantz, R. L. (1963). Pattern vision in newborn infants. *Science*, *140*, 296–297.
- Frank, M. C., Amso, D., & Johnson, S. P. (2014). Visual search and attention to faces during early infancy. *Journal of Experimental Child Psychology*, *118*, 13–26.
- Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, *110*, 160–170.
- Frank, M. C., Vul, E., & Saxe, R. (2012). Measuring the development of social attention using free-viewing. *Infancy*, *17*, 355–375.
- Gliga, T., Elsabbagh, M., Andravizou, A., & Johnson, M. (2009). Faces attract infants' attention to complex displays. *Infancy*, *14*, 550–562.
- Gluckman, M., & Johnson, S. P. (2013). Attentional capture by social stimuli in young infants. *Frontiers in Psychology*, *4*, 1–7.
- Gredebäck, G. & Daum, M. M. (2015). The microstructure of action perception in infancy: Decomposing the temporal structure of social information processing. *Child development perspectives*, *9*, 79–83.
- Gredebäck, G., Johnson, S., & von Hofsten, C. (2010). Eye tracking in infancy research. *Developmental Neuropsychology*, *35*, 1–19.
- Hayne, H. (2004). Infant memory development: Implications for childhood amnesia. *Developmental Review*, *24*(1), 33-73.

- Jankowski, J. J., Rose, S. A., & Feldman, J. F. (2001). Modifying the distribution of attention in infants. *Child Development, 72*, 339–351.
- Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40*, 1–19.
- Kolling, T., Óturai, G., & Knopf, M. (2014). Is selective attention the basis for selective imitation in infants? An eye-tracking study of deferred imitation with 12-month-olds. *Journal of Experimental Child Psychology, 124*, 18–35.
- Kwon, M.-K., Setoodehnia, M., Baek, J., Luck, S. J., & Oakes, L. M. (2014). Developmental Psychology The development of visual search in infancy : Attention to faces versus salience. *Developmental Psychology, 52*, 537–555.
- Leppänen, J. M. (2016). Using eye tracking to understand infants' attentional bias for faces. *Child Development Perspectives, 10*, 161–165.
- Libertus, K., & Needham, A. (2011). Reaching experience increases face preference in 3-month-old infants. *Developmental Science, 14*, 1355–1364.
- Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A.V., & Parlade, M.V. (2007) Individual differences and the development of joint attention in infancy. *Child Development, 78*, 938-954.
- Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. (2009). Foundations for a new science of learning. *Science, 325*, 284–288.
- Oakes, L. M. (2012). Advances in eye tracking in infancy research. *Infancy, 17*, 1–8.
- Óturai, G., Kolling, T., & Knopf, M. (2013). Relations between 18-month-olds' gaze pattern and target action performance: A deferred imitation study with eye tracking. *Infant Behavior and Development, 36*, 736–748.

- Papageorgiou, K. A., Smith, T. J., Wu, R., Johnson, M. H., Kirkham, N. Z., and Ronald, A. (2014). Individual differences in infant fixation duration relate to attention and behavioral control in childhood. *Psychological Science, 25*, 1371–1379.
- Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2004). Infant visual recognition memory. *Developmental Review, 24*, 74–100.
- Stoesz, B. M., & Jakobson, L. S. (2014). Developmental changes in attention to faces and bodies in static and dynamic scenes. *Frontiers in Psychology, 5*, 1–9.
- Taylor, G., & Herbert, J. S. (2013). Eye tracking infants: Investigating the role of attention during learning on recognition memory. *Scandinavian Journal of Psychology, 54*, 14–19.
- Taylor, G., & Herbert, J. S. (2014). Infant and adult visual attention during an imitation demonstration. *Developmental Psychobiology, 56*, 770–782.
- Wass, S. V., Forssman, L., & Leppänen, J. (2014). Robustness and precision: How data quality may influence key dependent variables in infant eye-tracker analyses. *Infancy, 19*, 427–460.

Table 1

*Mean peak fixation duration to the face and action area during the greeting phase, measured in seconds.*

Age	Peak fixation in area				pairwise t-test		
	Face		Action		<i>t</i>	<i>df</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
12 months	1.57	0.64	0.90	0.29	7.35	53	<.001
16 months	1.72	0.65	1.09	0.25	4.80	28	<.001

Table 2

*Descriptive statistics for the face-to-action ratio during the greeting and each demonstration segment of the videos for infants using shorter or longer fixation duration separated by age. Infants are first separated by peak fixation in the face area and then by peak fixation in the action area.*

	Face-to-action ratio							
	Greeting		1st demonstration		2nd demonstration		3rd demonstration	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Peak fixation in the face area								
Above median	.56	.016	.41	.016	.42	.019	.54	.019
Below median	.44	.016	.36	.015	.39	.019	.48	.019
Peak fixation in the action area								
Above median	.46	.017	.33	.014	.36	.018	.47	.019
Below median	.54	.017	.43	.014	.46	.018	.54	.019

**Highlights**

- Infant looking was tracked while a presenter repeated actions with objects
- 12- and 16-month-olds' attended more to the action during the first demonstration
- Infants increased attention to the presenter's face as the actions were repeated
- Prior object familiarity, but not presenter familiarity, influenced looking patterns

**Figure Captions**

*Figure 1:* Stimulus video after the mitten was removed from the puppet's arm and just before the presenter shakes the mitten to sound a bell inside (Note: Areas of interest are highlighted for coding purposes only and were not visible to the infant).

*Figure 2.* Face-to-action ratio for 12- and 16-month old infants for each video segment. Reference line at 0.5 indicates infants attend to face area as much as to action area. Error bars indicate 95% confidence interval.

**Figure 1**



Figure 2

