

The Development of Contour Processing Abilities in the Second Year

Samuel H. Forbes*

University of East Anglia

Kim Plunkett

University of Oxford

Abstract

Recent years have seen a rise in popularity of eye-tracking methods to evaluate infant and toddler interpretation of visual stimuli. The application of these methods make it increasingly important to understand the development of infant sensitivity to the perceptual properties implicated in such methods. In light of recent studies which demonstrate the use of pseudoisochromatic plates in testing infants for colour vision, we investigate the perceptual contouring abilities required to pass a colour vision test of this type. 115 (51 female) 16- and 19-month-old UK-based participants from the Oxfordshire region participated in this study. The evidence collected in this study indicated their ability to systematically fixate a contoured target, but that the speed at which they did so was much slower in the younger age group. These findings suggest the perceptual contouring abilities implicated in this study are still under development in the second year of life, and as such the results suggest a lower age limit for colour vision tests displayed in this format.

Keywords: Contour processing; Infant vision; Perception; Eye-tracking

*Contact author: samuel.forbes@uea.ac.uk. The authors are grateful to Irina Lepădatu and the Oxford BabyLab team for assistance with recruitment, and Karl Smith Byrne for his thoughtful input on the analyses. The manuscript also benefited immensely from the contributions of the anonymous peer-reviewers. The authors declare no conflicts of interest. This work was funded by a Rhodes Scholarship awarded to SHF. All data and analysis code required to analyse this dataset are available at https://osf.io/j8657/?view_only=64702cf578ce4ab89e9d884759281a87.

Introduction

In recent years, eye-tracking has become a common experimental method in developmental psychology. Many common iterations of this paradigm involve infants exploring two or more objects on a screen, while eye fixations are recorded. However, these methods, such as the intermodal preferential looking paradigm (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987), are contingent on a variety of participants' abilities, such as the ability to infer the boundaries of the objects from the background. Although a great deal is known about the development of infant visual perception in terms of their colour perception and visual acuity, far less is known about the developmental timeline for contour perception, that is the ability to infer a boundary of incomplete objects – the focus of the present study.

The basics of infant colour vision are largely in place by around 4 months of age (Bornstein, 1985; Franklin & Davies, 2004; Peeples & Teller, 1970, 1978; Teller, 1998; Teller, Civan, & Bronson-Castain, 2004), and develop over time (Knoblauch, Vital-Durand, & Barbur, 2001). Other features of infant visual perception, such as facial recognition and spatial frequency, develop at varying rates throughout childhood (Atkinson, Braddick, & Moar, 1977; Atkinson, Braddick, & French, 1979; Atkinson & Braddick, 1998; Atkinson, 2002; Braddick & Atkinson, 2011; Harris, Atkinson, & Braddick, 1976; Patel, Maurer, & Lewis, 2010). However, central to the ability to perceive and process objects and faces in the natural world is the ability to detect a contour (Gerhardstein, Tse, Dickerson, Hipp, & Moser, 2012; Taylor, Hipp, Moser, Dickerson, & Gerhardstein, 2014). Only having determined an edge of an object from the background noise around it can the object itself be recognised. This perceptual ability, a key to object recognition, is particularly difficult when the contour is incomplete. While an adult can generally infer these

boundaries (Kanizsa, 1987), this is considerably more difficult for young children.

While contour processing abilities are well-determined in adulthood, only a handful of studies have examined this in a developmental context. Burkhalter, Bernardo, and Charles (1993) demonstrated that while much of the overall structure of V1 is in place by the time of birth, the horizontal connections within the visual cortex develop slowly into an adult-like state over the first 24 months after birth, suggesting a slow developmental trajectory. These early results have been supported by behavioural studies investigating the thresholds at which young children can discriminate a contour from the background noise, which have found that while young infants can perceive contours, their ability to do so is far from developed (Gerhardstein, Kovacs, Ditre, & Feher, 2004). Thus from the points of view of both neural organisation and visual discrimination from background noise, contour perception is an ability which develops slowly.

One type of task that intrinsically requires more advanced contour processing abilities due to the inferred boundaries of objects, are colour vision tests such as the Ishihara Test (Ishihara, 1917), using pseudoisochromatic plates. In an Ishihara test, a shape or number is hidden in an image containing circles of varying sizes, and the image is only discriminable by contrast along a confusion axis. In a typical Ishihara test, the plates vary in their colours in order to test different colour deficiencies. However, there are two main problems involved with the implementation of such a test for infants. First, many of these tasks require a response, asking the participant to say what they saw, or to at least trace a figure of what they see; an infant is highly unlikely to be able to complete this task before three years of age. Second, the implementation of such a test is contingent on infants being able to recognise images presented in this format, and being able to perceive the boundaries of the objects. In order to complete a task using pseudoisochro-

matic plates, infants need to be able to fill in the outline of the image, a test of their contour processing abilities.

The modified pseudoisochromatic test has been demonstrated to be efficient in testing adults and children (Mercer, Drodge, Courage, & Adams, 2014; Pease & Allen, 1988). Pease and Allen (1988) developed testing plates that were rectangular in shape, and on one side had a target that differed along the relevant confusion axis (red/green or blue/yellow) from the background. Using a version of the Pease and Allen (1988) stimuli modified for infant use, Mercer et al. (2014) tested the colour vision of two groups of participants ranging from 3 to 23 months of age, along with an adult comparison group. They had an observer hold the plate, and observe head and eye motions of the participant through a peephole in the plate. The observer would rotate the plate a number of times until it was felt the the infant had a preference for one side over the other. The number of “trials” seen by each participant thus varied on the consistency of their individual looking behaviour. Mercer et al. found that a substantial proportion of infants fail the test up to around 16 months of age on the red/green plate, but from 17 months of age onward, these numbers decrease to around 10 percent of participants who fail the test. Interestingly this pattern of results does not occur for the blue/yellow plate, which was successfully completed by all but one participant.

The Mercer et al. (2014) study served as an important demonstration of colour vision testing in young infants, but it leaves the contouring aspect unanswered. When an adult completes an Ishihara test, they are often asked to state what they see, whereas in the Mercer et al. design they can simply be drawn to the difference in colour, and are not required to determine a shape or outline from the background. Additionally the experiment can continue until the observer is satisfied that the infant has made a decision, meaning that a decision might be more

likely than no decision. This raises the question: when this test fails to be successfully completed by an infant participant, what does it demonstrate? The first option is that infant colour vision at these younger age groups is not sufficiently developed that they can discriminate between the perceptually close colours used in such a test. In many cases this is likely to be true (Brown & Lindsey, 2013), but may not be the only reason such a test might fail in younger participants. The alternative possibility is that toddlers or infants below 17 months may be unable to successfully determine the contour and shapes of the target shown, meaning the only draw is the difference in colour, which may not hold enough of the participants interest without them distinguishing a discernible shape. Under this explanation, the contour perception abilities of the participants while still not fully developed, may inhibit their ability to systematically demonstrate a preference for the target. This would suggest that a test might fail below 17 months of age due to the contouring abilities being insufficiently developed to distinguish the outline of the image presented in this way.

In the developmental psychology literature, a standard method for testing infant visual discrimination is the looking-while-listening paradigm (Fernald, Zangl, Portillo, & Marchman, 2008). In this paradigm, an infant is typically presented with two objects, one either side of the screen, and given an auditory stimulus which should prompt looking toward one side over another. Using this paradigm, with a distractor on one side and a target on the other, allows us to test whether the shapes presented in a standard Ishihara style test using pseudoisochromatic plates are distinguishable to young participants, or whether they are likely to simply be drawn in by a difference in colour that appears on one side. Thus in the current study we test infant contour perception by asking them to distinguish between two possible targets discriminable from each other only by their

shape, and measuring their looking while doing so. An infant who has recognised the named object should systematically look to the named target (Fernald et al., 2008), whereas one who cannot should demonstrate no real preference between the two sides.

Since a key question for this study is the developmental trajectory of contour perception, participants at two different ages, 16 and 19 months, were recruited for the task. The aim of the present study was to explore the development of contour processing at these ages, with two main questions to investigate: First, if stimuli such as those proposed below are able to be discriminated, participants should demonstrate systematic looking to the target; Second, to examine the developmental trend of looking behaviour in the task, and investigate how it changes across age groups. In the design of this study, it was anticipated that fixation to the target could be slower than many looking-while-listening tasks, given the difficulty of determining an object outline in this task. Experiment 2 was thus designed to act as a control for Experiment 1, using actual photos of the same named objects, ensuring that any differences found in fixation speed were a function of the incomplete object boundary found in pseudoisochromatic plates, and not simply due to the fixation speed of this set of participants or a property of this selection of objects. Due to the relative ease of Experiment 2, and the need to maintain participant attention throughout the task, a shorter trial length was used for Experiment 2.

Experiment 1

Methods

Participants. A total of 115 participants participated in this task, $N = 58$ 16-month-olds (mean age 16.22 months, S.D. 0.46 months), and $N = 57$ 19-month-

olds (mean age 19.47 months, S.D. 0.66 months). All participants were recruited either online or at the local maternity ward. None of the participants demonstrated any signs of a serious vision deficiency, either to parents or to the researcher. The age groups were chosen to be as close together as possible, but one in the age group that had previously been shown to be unreliable with the Mercer et al. (2014) task, and one in the group that had been demonstrated to be reliably able to complete the task. An additional 5 (2 16-month-olds) participants were removed due to failure to calibrate, or due to no successful trials being completed.

Visual stimuli design. Visual stimuli were created based on a set of stimuli created by Aisch (2011). Stimuli consisted of common everyday images that would be familiar to young toddlers from everyday life (and were commonly known based on early Oxford Communicative Development Inventory estimates; Hamilton, Plunkett, & Schafer, 2000). These stimuli were presented in a format similar to the Ishihara plates. Unlike those created by Aisch, the current stimuli were instead coloured with hues as close as possible to the first Ishihara plate, that should be visible to anyone with better than monochromatic vision, in grey (CIE $L^*ab = 53.585, 0.000, 0.000$) and orange (CIE $L^*ab = 61.349, 35.270, 52.523$). Recolouring was performed in the Gnu Image Manipulation Program (www.gimp.org). Stimuli were designed so that the image would appear in orange, and the background would appear in grey, thus being visible to any individual with colour vision. 7 plates were created for this sample, containing an apple, a bottle, a butterfly, a cat, a dog, a duck and a tree. An example plate is included as Figure 1.

Plate selection. All seven plates were rated by 36 (20 female) adults through an online task. None of the adults reported any vision deficits. Those selecting the stimuli were asked to input as text what they thought the image represented in each plate was, and how clear that image was (on a scale of 0 - 100). 1

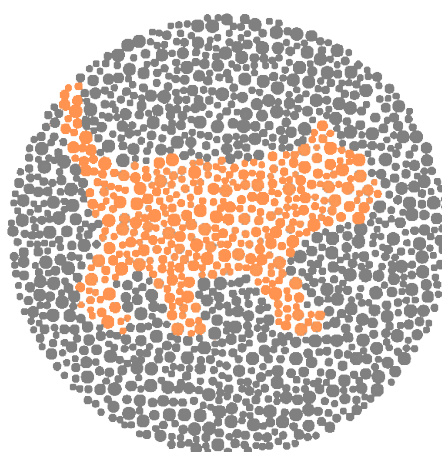


Figure 1. An example plate, depicting a cat, used in Experiment 1.

adult participant of those was excluded for misunderstanding the task, and only rating clarity up to 20. It was decided that any plate that was misidentified at least 25% of the time, or any plate with a clarity rating below 0.75 would be excluded. This 75-75 rule provided a simple heuristic for plate selection, and ensured that adult participants were satisfied that the plates were as designed, as well as being sufficiently clear. For each stimulus, over 80% of participants correctly identified the stimulus. For 4 of the plates (butterfly, dog, cat, and duck), all participants correctly identified the intended stimulus. However, the mean clarity rating of bottle was 0.68, meaning only the remaining 6 stimuli were included in the experiment. The ratings of all remaining stimuli are included in Table 1.

Table 1

Stimulus ratings for all stimuli included in the experiment. Identification is the proportion of adults who correctly identified the stimulus, while the clarity score is the rating out of 100, divided by 100.

Stimulus	Identification	Clarity
Apple	0.89	0.79
Butterfly	1	0.87
Cat	1	0.85
Dog	1	0.90
Duck	1	0.82
Tree	0.83	0.83

Auditory stimuli. All auditory stimuli were recorded by a female native speaker of Southern British English, speaking slowly and clearly, in an infant-directed manner. Stimuli consisted of sentences of the format: “*look at the xxx,*” where xxx was the name of each of the stimuli contained in the experiment. Recordings were cleaned and edited in audacity (www.audacityteam.com).

Procedure. Before arriving at the lab, participants’ caregivers were asked to fill out the Oxford Communicative Developmental Inventory (OCDI) (Hamilton et al., 2000) online. Participants arrived at the lab, and then had a short play session to acclimatise to surroundings, when consent forms were filled out by the caregiver. Participants were then seated on their caregivers laps, roughly 75cm from an eye-tracker and associated video screen. A nine-point calibration sequence was performed until at least 7 of the calibration points were calibrated correctly, after which the experimental trials commenced. In each trial, participants saw two stimuli, presented one on either side of the screen, on a white background, and were prompted to look at one by the auditory stimulus. In each trial, participants saw an attractive attention-getter for 2 seconds, after which point the stimuli immediately appeared on the screen. Following that, the auditory stimulus began, such that the onset of the target word occurred exactly 2 seconds after the visual stimuli onset. The trial continued for another 5 seconds after target word onset. Each participant saw a total of 6 experimental trials; each participant was randomly assigned to one of 5 experimental lists, that differed in their pairings of each stimulus. Trials were counterbalanced within subjects, such that each pair appeared twice, with each object in the pairing appearing as target and distractor. The time-course of a sample trial can be seen in Figure 2. Trials were presented using a custom script in MATLAB, and infant gaze data was recorded using a Tobii TX300 eye-tracker, recording at 120 Hz. The study was approved under the

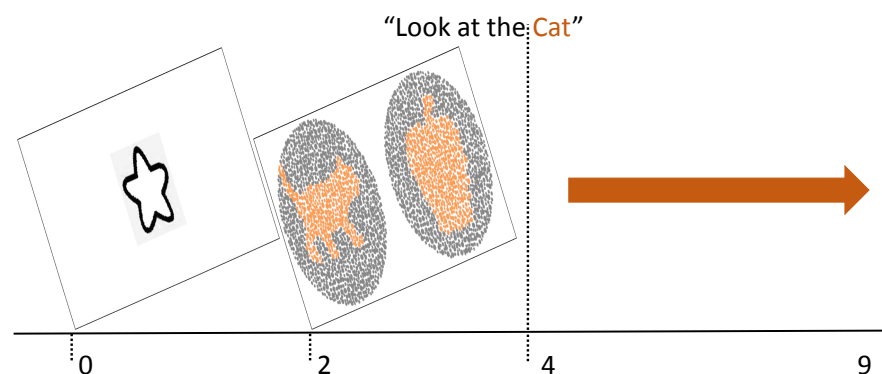


Figure 2. The time-course of a typical trial in Experiment 1.

University of Oxford Medical Sciences Interdivisional Research Ethics Committee, MS-IDREC- C1-2015- 071 (project title: Adjective and object property comprehension in children aged 3-36 months).

Analysis. Fixations for each participant were extracted using a custom MATLAB script, where a fixation was defined as stable gaze in one area of the screen for a minimum of 100ms, allowing for a certain amount of instability in the infant gaze. Data was cleaned so that only trials where participants looked to either of the images for 75% of the total trial time were included. The high threshold allowed flexibility due to the discrepancy in trial length between Experiment 1 and Experiment 2, as it is a threshold that can be applied equally across both experiments for direct comparison, and also allowed for the fact that 5 seconds after target naming is a relatively long time to keep interest. Only trials where parents indicated that the participant comprehended the label of the target object according to the OCDI were included in the analysis. This exclusion, as well as a number of participants who did not fill out the OCDI at all, meant that the final analysis included data from $N = 88$ participants (45 16 month-olds).

Data was then analysed with mixed models in R version 4.0.0 (R Core Team, 2020), using the `brm` function from the R package `brms` version 2.12 (Bürkner, 2017, 2018), and `eyetrackingR` version 0.1.8 (Dink & Ferguson, 2018). The Bayesian approach was chosen here to allow for a maximal random-effects structure that may not otherwise be possible with frequentist statistics due to the difficulty of convergence of generalised linear mixed-effects models with complex random effects structures. The Bayesian approach also allows for flexibility in modelling complex phenomena. Where the Bayesian approach differs from the more commonly used frequentist models, is that rather than estimating a single, maximum-likelihood estimate, it instead samples from all the model estimates that are consistent with the data and the prior information.

In both experiments, looking data were analysed using growth curve analysis (Mirman, 2014). For the growth curve analyses, fixations were first aggregated for each participant and each 100ms time bin (across trials), so that the amount of looks to the target and to the distractor were compared in each 100ms timebin, giving us a proportion of looks to the target. Growth curve analysis allows us to see how the probability of fixating the target over the distractor changes over time. These growth curves were in each case below modelled using a Bayesian logistic generalised linear mixed-effect model. For all models convergence was assessed with \hat{R} less than 1.1 for all parameters, and effective samples greater than 400, as well as no divergent transitions.

Results

The overall raw looking proportions over time can be seen in Figure 3. The points and standard errors in green indicate looking proportions from Experiment 1 (those in purple indicate performance in Experiment 2, discussed below). As we

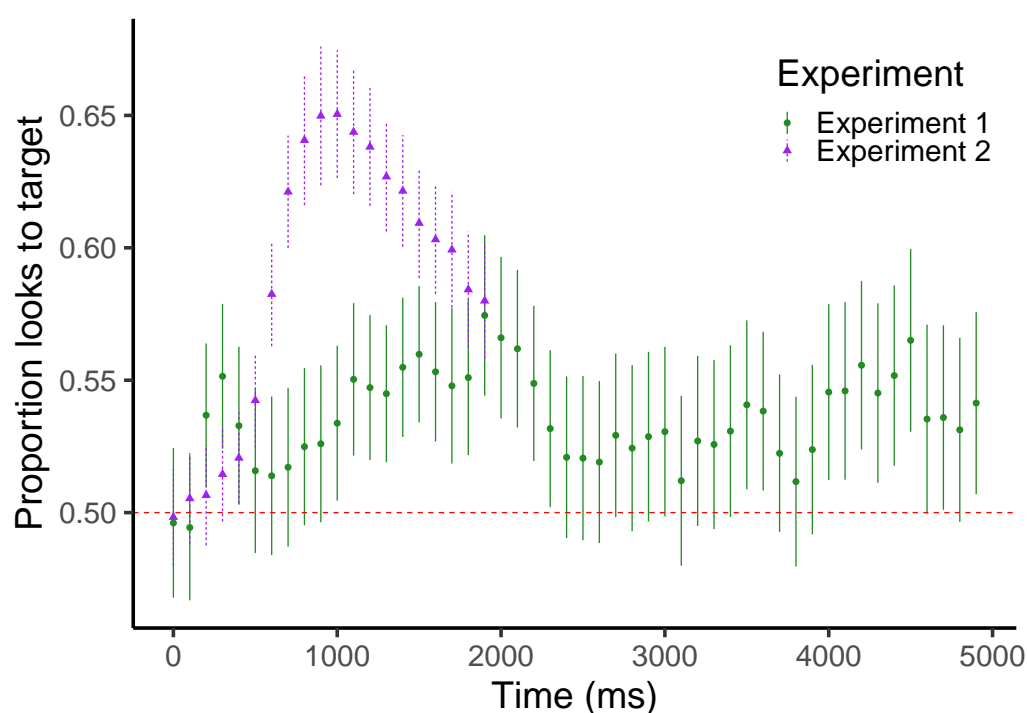


Figure 3. Looking proportions over time in Experiment 1 (green) and Experiment 2 (purple). Error bars indicate standard errors of the mean. The red dotted line indicates chance looking.

are only interested in looking to the target area of interest (AOI) vs the distractor AOI, proportions can be expected to be around chance (0.5) at target word onset, and then systematic looking should begin from around 300ms after target word onset (e.g. Bergelson & Aslin, 2017), as can be seen here. As can be seen from the figure, throughout the trial in Experiment 1, the probability of fixating the target is generally slightly above chance, reaching a peak about 2000ms after target word onset, when on average participants are fixating the target around 60% of the time.

Looking behaviour after the first three seconds was deemed prior to the analysis and data visualisation to be of little theoretical interest, as the effect of the target word will have diminished by that point (for similar arguments see e.g. Forbes & Plunkett, 2019a, 2019b). Thus the proportion of looks to the target dur-

ing the first three seconds of the trial following target word onset was modelled with growth curve analysis. As the outcome variable was the proportion of looking to the target, this required a binomial generalised linear model with a logit link function. We modelled time using cubic orthogonal polynomials to capture the curves within the data, so in concrete terms, our model included a linear time component, a quadratic time component, and a cubic time component in the fixed effects. The time terms were orthogonal (scaled and centered) so that they would not correlate with one another. To allow for the fact that older participants might have different looking patterns to younger participants, we also included a fixed effect of age group (16 or 19 months), which was difference-coded so that a positive fixed effect of age would indicate more looking by the older age group.

In order to allow for each participant having different patterns of looking, the slopes of the three time components discussed above were allowed to vary by subject. This allows us to fully estimate the time-course of looking to the target for each child in the study. Bayesian analyses allow us to use prior information to aid in estimation, and in this case we chose priors that are weakly informative, and centered around zero, so that they would not bias our results, but could easily be overcome by the data. Priors for the fixed effects were chosen to be Student's t distributions with a mean of 0 and a standard deviation of 2.5, and 3 degrees of freedom (similar to but slightly more flexible than the Cauchy priors previously recommended for logistic models, Gelman, Jakulin, Pittau, & Su, 2008; Ghosh, Li, & Mitra, 2018). For all analyses, priors for non-fixed effects were the RStan defaults (version 2.19.1 Stan Development Team, 2016). We set the model to sample for 2000 iterations, of which 1000 was a warmup, and this was done on four chains, so we could be sure of the starting value not influencing the estimates.

The estimates from the fixed effects of the model can be seen in Table 2. In

general, we can see that if the 95% credible interval of any parameter excludes zero, there is reason to believe there is a reliable, non-zero effect. In this case, the model output indicates little evidence of a main effect of age, or of any of the polynomial time terms, where the 95% credible interval of each of the parameters, with the exception of the model intercept, includes zero.

Table 2

Population level effects of the model output of Experiment 1. Columns indicate estimated value, Estimated error, lower and upper 95% credible intervals, and the number of effective samples. Orth time refers to each of the orthogonal polynomial time terms, linear, quadratic, and cubic. Note that it is possible for the effective sample size to exceed the number of samples in the model under certain conditions.

	Estimate	Est. Err.	l-95% CI	u-95% CI	Eff. Samples
Intercept	0.42	0.19	0.06	0.79	1659
Orth. Linear Time	0.36	0.91	-1.45	2.18	1735
Orth. Quad. Time	-0.83	0.85	-2.51	0.85	1943
Orth. Cubic Time	0.27	0.42	-0.54	1.08	2298
Age	0.52	0.36	-0.16	1.23	1698
Orth. Linear Time:Age	1.74	1.61	-1.23	5.18	1966
Orth. Quad. Time:Age	0.39	1.42	-2.38	3.17	2210
Orth. Cubic Time:Age	-0.23	0.77	-1.75	1.27	2332

Figure 4 demonstrates the model fit to the data from Experiment 1 for the the 16- and 19-month-olds. From examining Figure 4, it can be seen that while the 19-month-old participants tend to fixate the target objects around the 1000ms mark, the peak in target looking for the 16-month-olds is much later, around the 2000ms mark. The time to peak fixation is late compared to what is normally seen from a target-present intermodal preferential looking task, where peaks in fixation to a target should typically appear before around 1000ms at these ages. The ribbons indicating the 95% credible interval of the model estimate are also a useful heuristic for visualising when looking was systematically above chance¹,

¹It is important to note that this is not always the case, and that the reliability is also dependant on the random effect structure, but it is a useful identifier in this kind of data visualisation, one that should not be taken as a concrete rule.

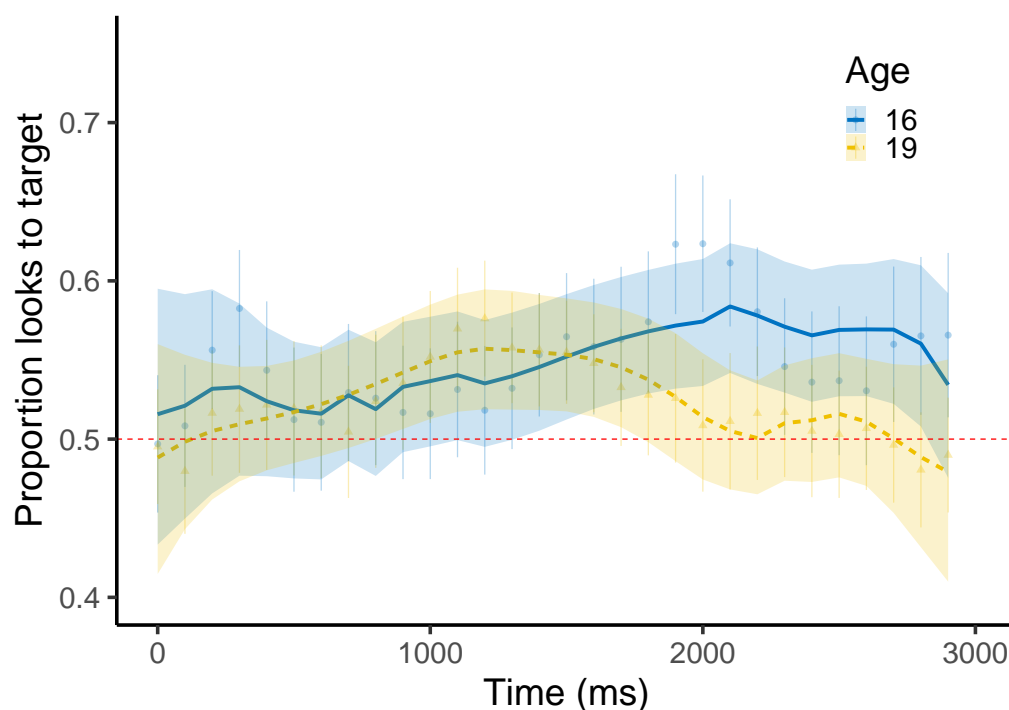


Figure 4. Model fit to the data from Experiment 1. Points with error bars indicate mean and standard errors of the data, lines and ribbons indicate the model estimate and the 95% credible interval of the estimate.

based on visualising when the entirety of the ribbon is above the dotted line which otherwise indicates chance looking. Both groups of participants appear to systematically fixate the target based on Figure 4, with periods of reliable looking above chance, demonstrating an ability to complete the task.

In order to ascertain whether there the looking patterns to the target differed by age group, the estimates from the posterior samples from each age group were extracted from the model using the tidybayes package version 2.1.1 (Kay, 2020), and at each time point, the mean difference (19 month old minus 16 month old) was calculated, along with the maximum difference and minimum difference. These can be seen in Figure 5. For the first two seconds the difference between the two age groups centres around zero indicating no difference between the groups,

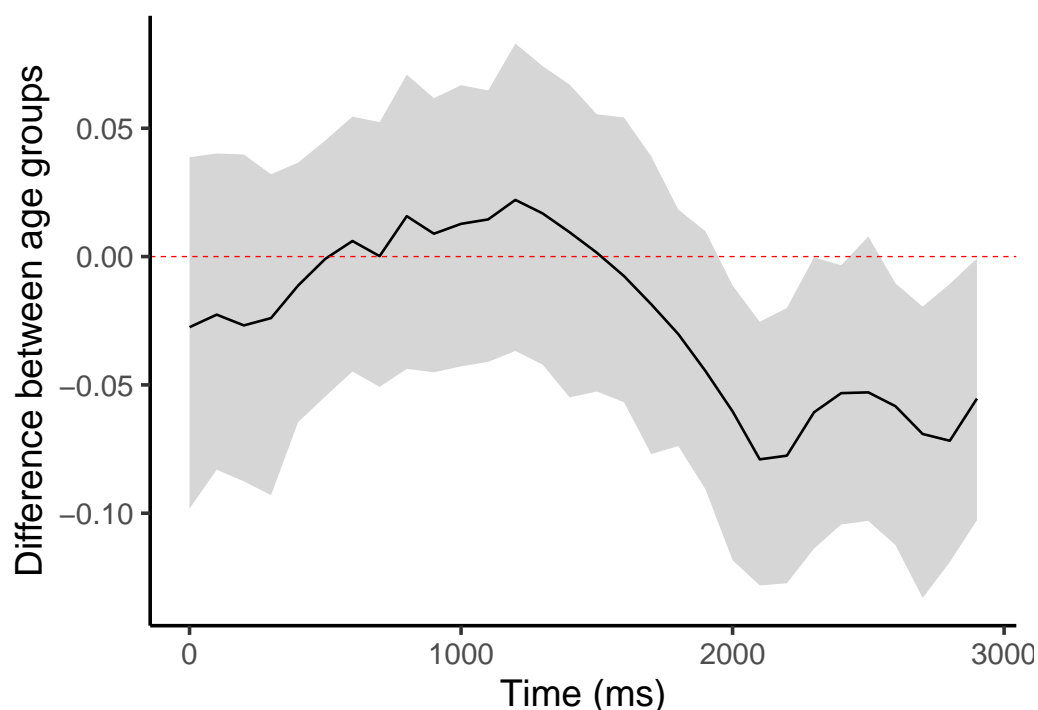


Figure 5. Differences between the two age (19m.o. - 16m.o.) groups in Experiment 1 based on the posterior sampling of the data. The thick black line indicates the mean difference, while the ribbon indicates the minimum to maximum difference. The curves can be considered different where the entirety of the interval is above or below zero.

but after two seconds the mean and maximum of the difference drops below zero, indicating that from two to three seconds after target word onset, the 16-month-olds are looking to the target reliably more than the 19-month-olds. This is indicative of the fact that the 19-month-olds look to the target earlier, then lose interest, whereas the 16-month-olds are slower to move, and so their peak in looking to the target comes after the older participants have fixated the target and looked away.

The results above indicate that both age groups can successfully locate the target image. However, the younger age group do so later than the older age group. This suggests two distinct possibilities: first, that the difference is symptomatic of a general difference in processing and fixation speed between the two age groups

(Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998), or second, that the difference is caused by difficulties specific to the stimuli utilised in Experiment 1. Experiment 2 examines these questions under similar conditions to those employed in Experiment 1, but by manipulating the stimuli so they are instead just pictures of the real objects as photographs, allowing us to test whether this difference in processing speed is normal between the two age groups.

Experiment 2

Methods

Participants. A subset of the participants who completed Experiment 1 ($N = 31$ 19-month-olds) and ($N = 30$ 16-month-olds) also completed a second experiment. These participants were those who were willing to continue after the first task, and excluded the first handful of participants, who completed Experiment 1 prior to the creation of the task used in Experiment 2.

Stimuli. The plates used in Experiment 1 were replaced with photos of each of the objects used in the experiment.

Procedure. The procedure for Experiment 2 was almost identical to that of Experiment 1, but the participants who partook in the experiments were given a short break before taking part in Experiment 2. It was imperative that Experiment 2 was performed after Experiment 1, not before or interleaved, because having been exposed to typical exemplars of the images (e.g. a normal photo of a cat) it would make it much easier for them to find the correct image in Experiment 1. Trials for this experiment were cut shorter to allow for more trials, leaving only 2 seconds after target word onset, as it was expected prior to the creation of the task that participants would have less difficulty in locating the target image when presented with real photographs, and that if the task were too simple and slow,

it may lead to participants fussing out. Participants saw 30 trials in a randomised order in this experiment for improved statistical power (due to the smaller number of participants compared to Experiment 1); due to the increased number of trials, all participants saw the same experimental list, which paired each object with each other object.

Analysis. Data for this experiment was analysed in an identical fashion to that of Experiment 1. After excluding participants who did not know the target words according to the OCDI, a total of 47 participants (27 16-month-olds) remained.

Results

The raw looking data for Experiment 2 can be seen from the purple curve in Figure 3. In contrast to Experiment 1, the participants in Experiment 2 show a steep and rapid increase in target looking following the target word onset, peaking around 1000ms, where they look to the target around 65% of the time.

As in Experiment 1, we used growth curve analysis to estimate the probability of looking to the target over time. As such, a similar model was fit to the one specified in the previous experiment. Once again we used a logistic growth curve analysis as we were estimating proportions, and we again used cubic orthogonal polynomials of time to estimate the shape. The key difference between this model and the one in Experiment 1 is that the trial is only 2000ms long after target word onset, rather than the 3000ms used above.

Table 3 highlights the model output from the fixed effects. The 95% credible interval for age intersects with zero, but of particular interest is the fact that the interaction between age and the quadratic time term (ot^2) does not intersect with zero, indicating a difference in time-course between the two age groups. The

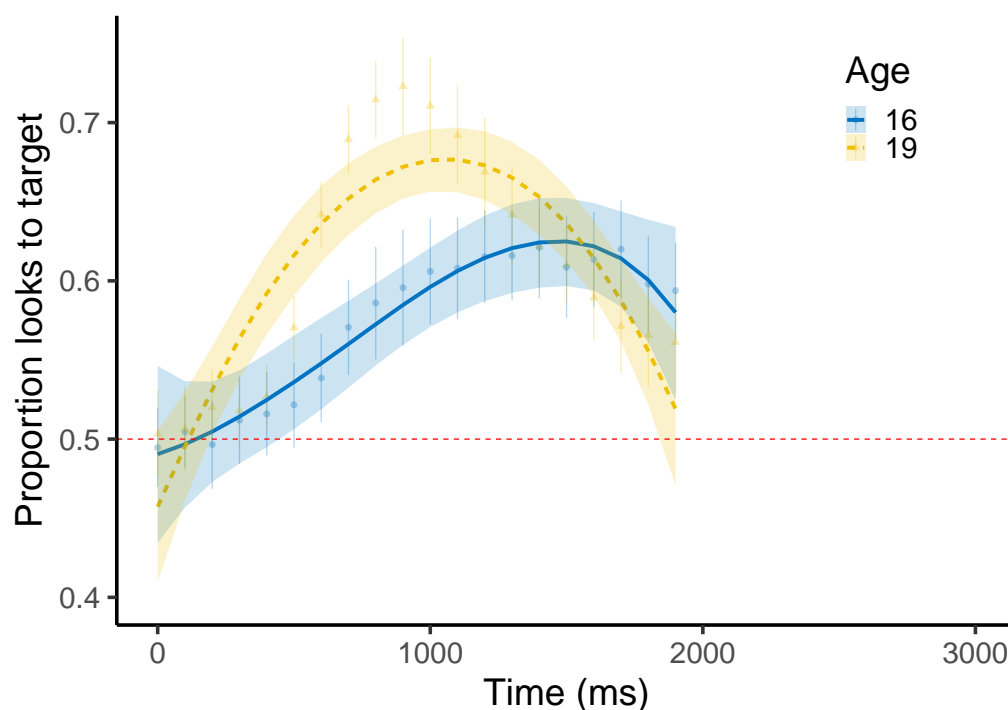


Figure 6. Model fit to the data from Experiment 2. Points with error bars indicate mean and standard errors of the data, lines and ribbons indicate the model estimate and the 95% credible interval of the estimate.

model fit to the data can be seen in Figure 6. The peak looking period for the 19 months participant group is slightly earlier (1000ms vs. 1500ms) than in the previous experiment, reflecting the decreased difficulty participants in this age group have when visually searching for the target when it is a picture, rather than presented as in Experiment 1. Overall, the proportions of looking to the target are also much more obviously above chance in Experiment 2 than in Experiment 1 for this age group, demonstrated by the lower 95%CI being above chance for most of the trial in both age groups, although it is important to remember that not all of the participants who completed Experiment 1 also did Experiment 2, so due to the differing samples this is not a direct comparison between the two groups.

Unlike the small difference between the peaks in target looking between Ex-

Table 3

Population level effects of the model output of Experiment 2. Columns indicate estimated value, Estimated error, lower and upper 95% credible intervals, and the number of effective samples. Orth time refers to each of the orthogonal polynomial time terms, linear, quadratic, and cubic. Note that it is possible for the effective sample size to exceed the number of samples in the model under certain conditions.

	Estimate	Est. Err.	l-95% CI	u-95% CI	Eff. Samples
Intercept	0.39	0.07	0.24	0.54	1354
Orth. Linear Time	0.67	0.21	0.27	1.1	1851
Orth. Quad. Time	-0.75	0.18	-1.1	-0.39	2039
Orth. Cubic Time	-0.14	0.09	-0.32	0.05	2474
Age	-0.15	0.15	-0.44	0.14	1450
Orth. Linear Time:Age	0.43	0.43	-0.43	1.28	1281
Orth. Quad. Time:Age	0.88	0.36	0.19	1.57	2210
Orth. Cubic Time:Age	-0.13	0.19	-0.49	0.26	2160

periment 1 and 2, Figure 6 demonstrates that at 16 months there is a considerable difference in looking patterns between the first and second experiments. In Experiment 2, the peak in looking occurs over 1000ms earlier than the peak in Experiment 1, with also much less of a difference between the two age groups, suggesting that for 16 month-olds, locating the target was much more difficult in Experiment 1 than in Experiment 2. It is additionally worth noting that the cubic polynomial is unable to fully capture the rapid rise in looks to target around the 1000ms mark for the 19 month-olds, due to the conservative nature of the polynomial. Despite this, the model is able to capture the difference between the two age groups. While there is a difference at 19 months between the two experiments, it is not as dramatic as the difference appears to be at 16 months. While there is a difference in time to fixate the target between Experiments 1 and 2, such conclusions cannot be drawn without directly comparing the same participants in each condition. Below, we make direct comparisons between the two experiments to further explore these differences.

Combined Results from Experiment 1 and 2

Timecourse analysis

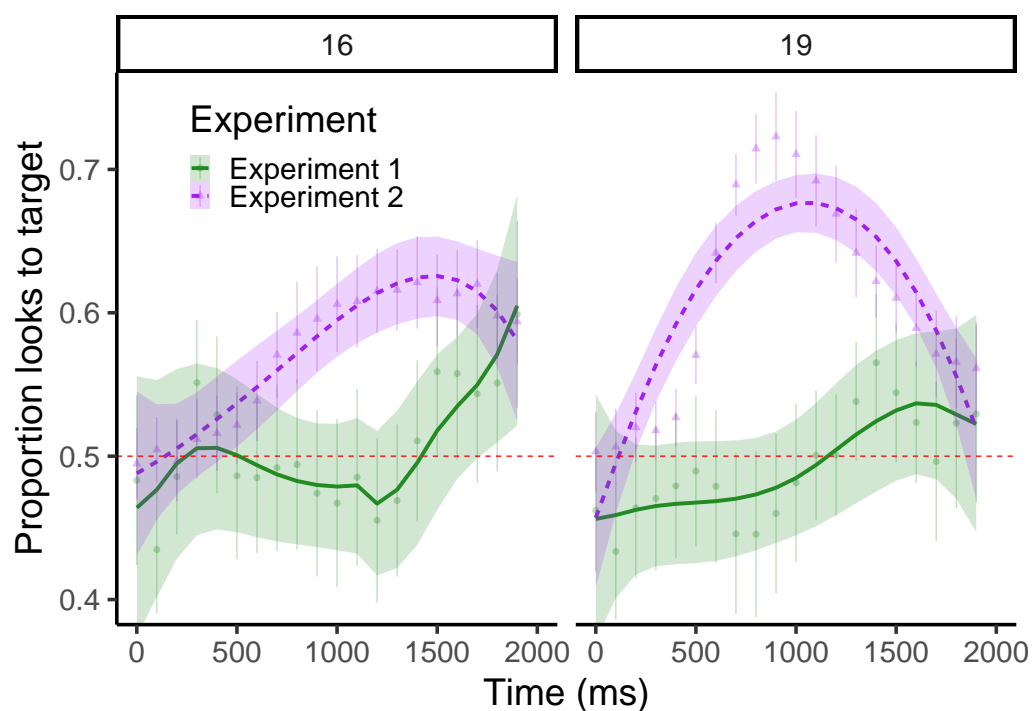


Figure 7. Model fit to the combined data of participants who completed both Experiment 1 (green), and Experiment 2 (purple). Points with error bars indicate mean and standard errors of the data, lines and ribbons indicate the model estimate and the 95% credible interval of the estimate.

For the below analyses, only participants who completed both Experiments 1 and 2 were included (47 participants). First the data was combined in order to estimate the differences between conditions using growth curve analysis, and participants who did not complete both experiments were excluded. As above, we used a logistic model to estimate the proportion, and modelled time with cubic orthogonal polynomials. However, in addition to the fixed effect of Age, which was also included in the previous models, we now add an additional predictor of condition (Experiment 1 vs Experiment 2), also difference coded such that a

positive effect of condition would indicate more looking in Experiment 2 than Experiment 1. Additionally, all interactions between fixed effects were also assessed. Because we aimed to calculate a growth curve for each participant in each condition, rather than have the time terms vary by participant as we did above, in this model we had time terms vary by an interaction between participant and condition. The time period modelled was the first 2000ms after the target word onset for both experimental conditions, to keep the time period consistent.

Table 4

Population level effects of the combined model output from participants in Exp1 and Exp2. Columns indicate estimated value, Estimated error, lower and upper 95% credible intervals, and the number of effective samples. Orth time refers to each of the orthogonal polynomial time terms, linear, quadratic, and cubic. Cond refers to the effect of experimental condition (experiment 1 vs 2). Note that it is possible for the effective sample size to exceed the number of samples in the model under certain conditions.

	Estimate	Est. Err.	l-95% CI	u-95% CI	Eff. Samples
Intercept	0.23	0.09	0.06	0.41	2834
Orth. Linear Time	0.86	0.31	0.25	1.47	2907
Orth. Quad. Time	-0.15	0.3	-0.73	0.45	3232
Orth. Cubic Time	0.09	0.13	-0.16	0.33	3747
Age	0.06	0.18	-0.28	0.41	2600
Condition	0.32	0.17	-0.02	0.67	2981
Orth. Linear Time:Age	0.46	0.61	-0.69	1.72	2547
Orth. Quad. Time:Age	0.74	0.58	-0.38	1.89	2756
Orth. Cubic Time:Age	0.37	0.24	-0.11	0.83	3787
Orth. Linear Time:Condition	-0.28	0.6	-1.46	0.88	2994
Orth. Quad. Time:Condition	-1.04	0.59	-2.17	0.16	3582
Orth. Cubic Time:Condition	-0.39	0.25	-0.87	0.08	3868
Age:Condition	-0.36	0.33	-0.99	0.28	2540
Orth. Linear Time:Age:Condition	0.09	1.09	-2.09	2.23	3329
Orth. Quad. Time:Age:Condition	0.31	1.09	-1.82	2.38	3424
Orth. Cubic Time:Age:Condition	-0.88	0.47	-1.8	0.04	4090

Figure 7 demonstrates the differences between the two experimental conditions. While the 2000ms window captures the peak looking in Experiment 2, shown in purple, the peak looking window is only starting to begin around the

2000ms mark for Experiment 1, in green, consistent with the looking patterns of the full sample size for Experiment 1. So from the estimated means shown in the graph we can see that while a typical 16 month-old participant would be looking around chance up until around 1500ms in Experiment 1, the same average participant is likely to be looking reliably to the target side from about 500ms onward in Experiment 2. Additionally, participants systematically fixate the target much more in Experiment 2 than in Experiment 1, demonstrating the additional difficulty presented by the stimuli in Experiment 1.

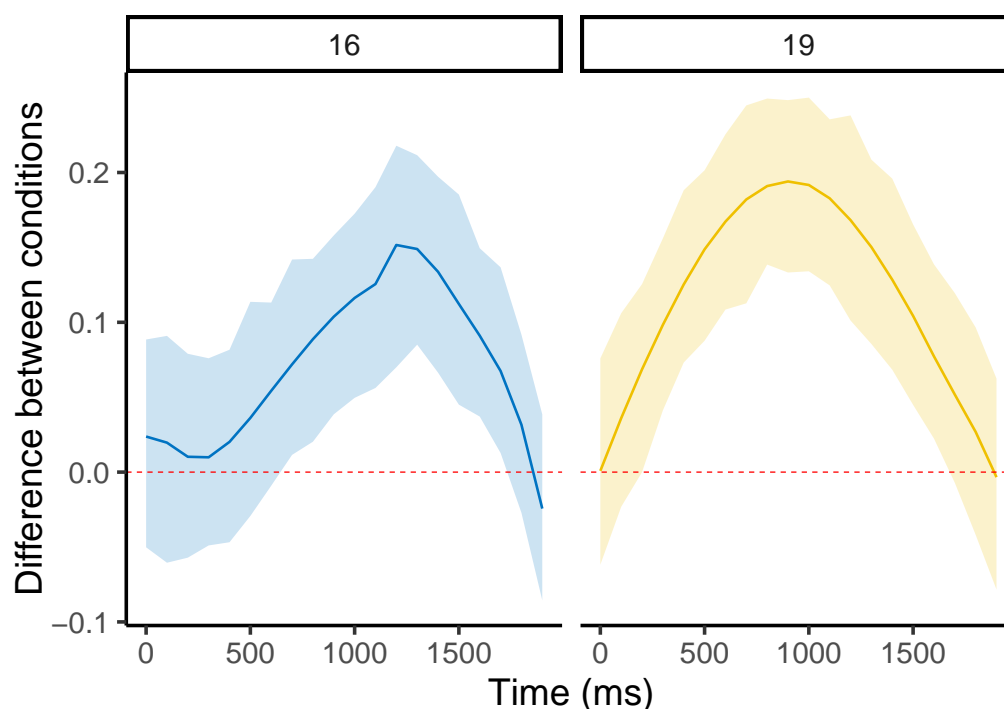


Figure 8. The difference between the two conditions estimated from the posterior samples of the model. A positive difference indicates more looking to the target in Experiment 2 than Experiment 1. Areas where the shaded area does not overlap with zero can be considered time periods where the two conditions diverge.

The results of the time course model are displayed in Table 4. Overall, with the exception of the linear time term (ot1) the credible interval of all effects contain zero. This may seem surprising as Figure 7 appears to show a strong difference

between conditions, but the presence of the polynomial time terms as well as the presence of condition interacting with participant in the random effect structure makes using the credible interval on its own from the credible interval extremely difficult.

However, our interest is in estimating divergences between conditions over time, so as we did with Experiment 1, we can use the posterior draws to assess whether at any point in time, the difference between conditions is reliably non-zero. The posterior draws were extracted from the model, and then the mean difference in conditions (Experiment 2 minus Experiment 1) was extracted for each timebin. The results of this can be seen in Figure 8. This demonstrates that for both age groups, participants are looking more to the target in Experiment 2 than Experiment 1 for most of the trial. This difference is slower to emerge in the younger age group, where the minimum of the difference begins to go above 0 around 700ms after target word onset. Thus we can see there is a large divergence between the two experiments in both age groups, although this is more pronounced in the 19-month-olds.

Switch analysis

In addition to the looking over time throughout the trial, the number of switches in looking between the two areas of interest (AOIs) was also analysed for the combined data only (similar analyses can be seen in Yurovsky & Frank, 2017; Fernald et al., 2008). The theory behind this was that participants who realise that the target word relates to the target should initially be more likely to switch to the target AOI from the distractor AOI, than the reverse. Likewise participants who relate the target word to the target should be **faster** to switch from the distractor to the target, than they are to switch from the target to the distractor. The switches

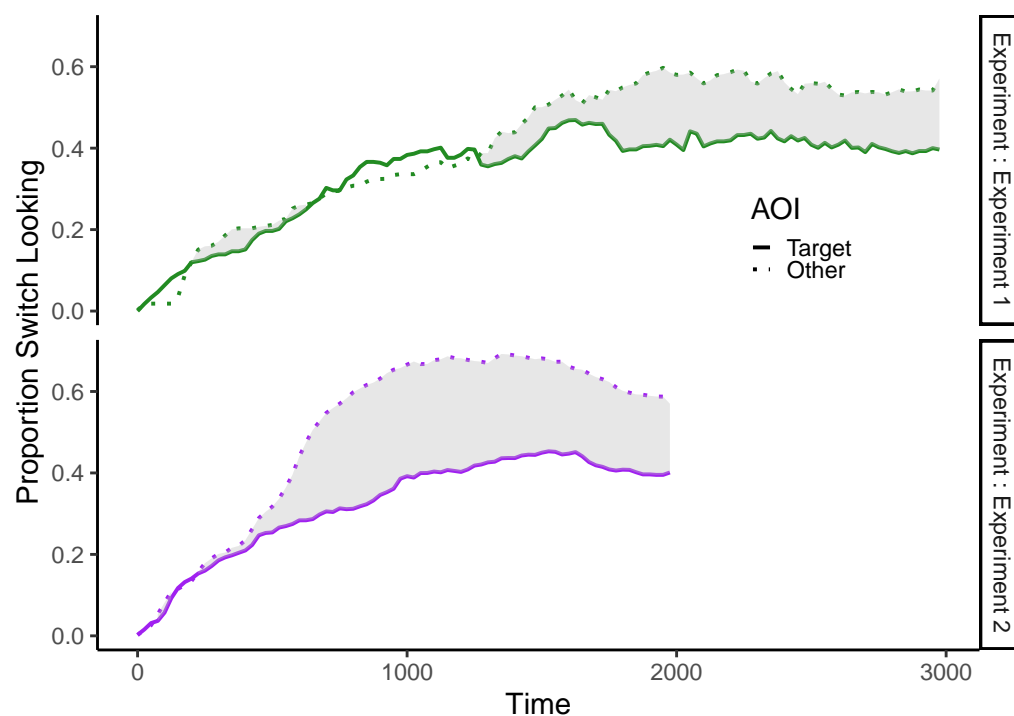


Figure 9. Switches in looking over time in Experiment 1 (top) and Experiment 2 (bottom). Shaded areas indicate where participants switched more from the distractor to the target, indicating recognition of the target. The solid line indicates the proportion of switches away from the target, and the dotted line indicates the proportion of switches to the target.

in looking to the target and distractor over time provide a useful inferential insight into the relative difficulty of the task (Yurovsky & Frank, 2017). Thus if a participant has understood that the auditory stimulus refers to the target object, they should make more switches to the target AOI than away from the target AOI.

Figure 9 indicates the proportion of children who switched to look at the other object from the object they started with. The solid lines indicate participants who began by looking to the target, whereas the dotted line indicates participants who began looking to the distractor. So as an example, in Experiment 2 (purple, lower panel), 1000ms after target word onset, around 30% of participants who began by looking to the target will have switched to the distractor. By contrast, at

the same point in time, over 60% of participants who were looking to the distractor had switched to look at the target. The shaded regions indicate periods of time where participants are more likely to switch to the target than away from it. In this sense, the larger the grey shaded area, the better the indication that participants are switching towards, rather than away from, the target.

While purely an visual tool, the figure indicates the participants switch more, and switch earlier, to the target than away from it in Experiment 2 (purple, bottom panel) than in Experiment 1 (green, top panel). The results demonstrate that participants switch from the distractor to the target more reliably before the 1000ms mark in Experiment 2 (bottom panel), whereas in Experiment 1 (top panel) this happens less obviously, and much later. The later recognition of the named target in Experiment 1 as opposed to Experiment 2, shown both in looking proportion and switch proportion, indicates a slower processing speed.

Onset-contingent switch times

Based on the above analysis, the latency of switching from the AOI that was fixated at target word onset was measured. The time to first switch from one AOI to the other after first fixation was calculated using eyetrackingR (Dink & Ferguson, 2018). The rationale for this analysis was to examine the switch time of participants to switch to the target as well as away from it, and how that effect is modulated by the relative difficulty in processing the task (Fernald et al., 2008). Thus if contour processing causes the stimuli in Experiment 1 to be harder to process than those of Experiment 2, this should be reflected by much slower switch times in Experiment 1 than in Experiment 2. Figure 10 demonstrates the switch times from the target AOI to the distractor AOI (grey), and switch times from the distractor to the target (red). Of particular note is the much faster switch times

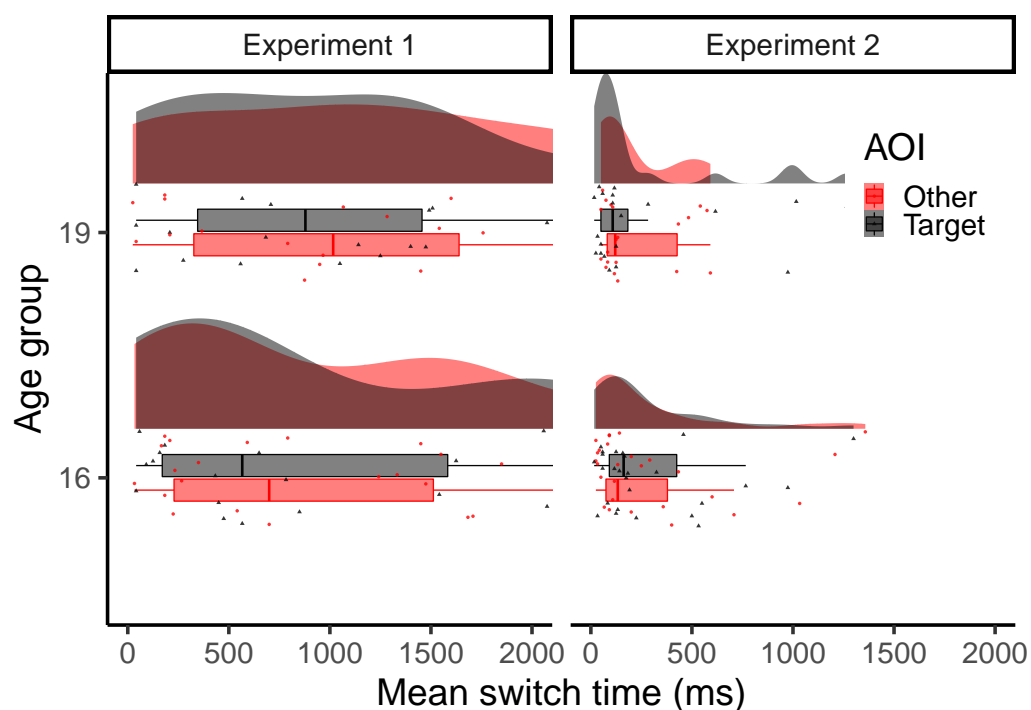


Figure 10. Raincloud plot of mean switch time in Experiment 1 (left) and Experiment 2 (right). Results from 19 month-olds are on the top, and 16 month-olds are on the bottom. In grey are switches from the target AOI to the non-target AOI, while the reverse is in red.

in Experiment 2, in the right panel, than in Experiment 1, on the left. The data is presented as a raincloud plot (Allen, Poggiali, Whitaker, Marshall, & Kievit, 2019). The top of each figure is a density plot indicating the density of points at that point in time. Below that is a box plot indicating the median and interquartile range of the data. The points indicate the raw data themselves.

Since switch times are not normally distributed, but are instead positive and skewed generally, these data were modelled with a model from the gamma family, using the `brm` function in `brms` (Bürkner, 2017, 2018) with a log link. For this model, the fixed effects were age (16 or 19 months) and condition (Experiment 1 or Experiment 2), as well as the AOI where the participant was looking at target word onset (target or distractor). A random intercept was added for each partici-

pant, and for each trial, to allow for patterns of response affected by participants and individual trials. On this occasion the prior for the intercept was a Cauchy distribution with a mean of zero and a scale parameter of ten, and for other fixed effects it was a Cauchy distribution with a mean of zero and a scale parameter of 2.5 (Gelman et al., 2008). Since the model is a gamma model, the prior for the shape parameter was selected to be the default gamma distribution in the package brms.

Table 5 contains the population-level effects of the model estimating the switch times. Of particular note is that the model output indicates strong evidence of an effect of condition, such that participants are switching much faster in Experiment 2 than in Experiment 1, demonstrated by the 95% credible interval not intersecting with 0. There is also evidence of an interaction between condition and age group. Participant switch looking speed was observed to be faster in the control condition in Experiment 2 than in the condition which tested contouring abilities, in Experiment 1.

Table 5

Estimated fixed effects of model on infants switch times. First AOI refers to the AOI at target word onset, and condition refers to the experimental condition (Experiment 1 or 2). Note that it is possible for the effective sample size to exceed the number of samples in the model under certain conditions.

	Estimate	Est. Err.	l-95% CI	u-95% CI	Eff. Samples
Intercept	6.82	0.03	6.75	6.89	4462
First AOI	0.04	0.06	-0.09	0.16	3172
Condition	0.55	0.07	0.42	0.68	4689
Age	-0.13	0.07	-0.27	0	3245
FirstAOI:Condition	-0.18	0.13	-0.44	0.08	3484
FirstAOI:Age	-0.21	0.13	-0.46	0.05	3167
Condition:Age	-0.31	0.13	-0.56	-0.06	3372
FirstAOI:Condition:Age	-0.04	0.25	-0.53	0.45	3257

Discussion

In the present study, two experiments were run to analyse the contouring abilities of toddlers to examine whether their developing contour processing enabled them to distinguish a named target when presented against a distractor in pseudoisochromatic plate format. In the first, infants were tested with modified testing plates similar in design but not colour or luminance to pseudoisochromatic tests of colour vision, depicting outlines of familiar images, and were prompted to look at one of them by an auditory stimulus. In the second, the plates were replaced with real images of the same objects. In both experiments participants demonstrated preferential fixation of the target, indicating object recognition, showing that in general they could successfully tell a target from a distractor by inferring the outline. However, while the 19-month-olds were slightly slower to fixate the target when it was presented in Experiment 1 as opposed to a photo of the image in Experiment 2, for the 16-month-olds, this difference in processing speed was much larger, with a peak fixation as late as 2000ms in Experiment 1, as opposed to around 100 in Experiment 2. These findings are confirmed in the combined analyses, where we see direct evidence of slower fixation and less fixation in Experiment 1 than in Experiment 2, even when including only those participants who completed both experiments. Likewise, the switch analysis indicates more switches and faster switches to the target than away from it in Experiment 2, where pictures of the real objects were used, than in Experiment 1 where contour processing was examined.

These results raise several important considerations. First, the results tentatively indicate that contour processing has developed sufficiently for these stimuli to be used to test for colour vision deficiencies at 16 months, but due to the slow

fixation speed, it may be at the limit of these abilities. However, if the additional difficulty of the more fine-grained colour perception were added to these plates to allow them to be used for a colour vision test, it is likely to raise the level of difficulty even higher. The results demonstrate that fixating an outline of an image presented in pseudoisochromatic plates is a difficult but not an insurmountable challenge to toddlers as young as 16 months. Additionally, presenting the task as an looking-while-listening task allows reliance on the participants ability to fixate the image, rather than describe the image, provided the object label is known to the participant.

The second consideration raised by the results of the present study is that while it appears that stimuli such as these can be processed by toddlers as young as 16 months, it is unlikely to work with age groups much *younger* than 16 months. While the 16-month-old age group were able to locate the target image, they did so much slower than the 19-month-old age group, and much slower than their own performance in Experiment 2. Additionally the wide credible interval supports that claim, demonstrating that in Experiment 1, participants were on average able to locate the target from the distractor, but there was a non-zero possibility of an average child not being able to systematically fixate the target. Put another way, while in general participants were able to systematically fixate the named target from 16 months of age, the credible interval for doing so was wide, indicating that this would not always be the case. This suggests that while the younger participants were able to fixate the target, they did so with a considerable degree of difficulty.

It is important to note that it is difficult to answer this question for infants much younger than 16 months using the present task, as a large proportion of participants would not know the object labels.

The difficulty experienced by the 16 month-olds in Experiment 1 is likely attributable to the stage of development of their contour processing abilities. Based on these findings, contouring is a perceptual property that develops slowly, much slower than, for example, colour perception (Bornstein, Kessen, & Weiskopf, 1976; Bornstein, 1985; Atkinson, 2002). A partially-occluded image or an image with an incomplete outline is likely to present varying levels of difficulty for the individual, depending on the developmental stage of the individual in question. An alternative conclusion is that participants at 16 months have a slower processing speed overall than the 19-month-olds (Fernald, Thorpe, & Marchman, 2010). Slower processing would interact with the additional difficulty imposed by the degraded image in Experiment 1 leading to the slower target recognition observed. However, the results of Experiment 2 suggest that this is not the case. The overall processing speed for 16-month-olds is not dramatically slower than for 19-month-olds as seen in Experiment 2, which indicates that the slower time to fixation seen in Experiment 1 is a function of the stimuli used. These findings are also in accordance with studies that have shown that the neural architecture for contour processing may develop slowly over the first two years of life (Burkhalter et al., 1993; Taylor et al., 2014).

A further consideration in interpreting the results of the present study is that rather than a difficulty inferring the outline of the image presented on the screen, the infant participants in the present study may simply see the image as an atypical member of that category. Under that explanation, infants would take longer to work out that the image of, for example the cat presented in this study, belongs to the category of cat, analogous to atypical exemplars when testing for object labels or object feature labels (e.g. Forbes & Plunkett, 2020; Horst & Samuelson, 2008; Meints, Plunkett, & Harris, 1999; Swingley, 2010; Wagner, Dobkins, &

Barner, 2013; Wagner, Jergens, & Barner, 2018). However, substantially younger participants than those included in the present study have successfully categorised non-photographic images. Chen and Westermann (2018) for example tested categorisation of 10 and 15 month-old participants with animated cartoon animals, and yet these images did not add additional difficulty to the study, despite only being outlines, and therefore atypical exemplars. While the images presented in the current study are harder again, a typically-presented outline should be sufficient for matching with a named label. Similarly, a variety of two-dimensional cartoon-style images have often been used to test categorisation abilities in young infants (Althaus & Westermann, 2016; Plunkett, Hu, & Cohen, 2008; Younger, 1985). Much younger infants have not incurred difficulties while using images such as those presented by Younger, so while there is no doubt that the images used in Experiment 1 are atypical examples of those objects, it should not reasonably be beyond the scope of their abilities.

It is important at this point to reconcile the differences in findings between the present study and that of Mercer et al. (2014), which showed that at least along the blue/yellow confusion axis, infants could locate a target in a pseudoisochromatic plate. In the Mercer et al. study, there was a target on only one side of the plate, so the only requirement is for the participant to distinguish between the colours located on the confusion axis. They were not required to necessarily identify a shape. By contrast, in the present study there was a target on both sides, and participants had to distinguish between the two by their shape. Additionally, in the present study participants had only one attempt to correctly distinguish the target from the distractor, as opposed to the Mercer et al. study where trials continued until a decision was reached.

In conclusion, the present study has examined the contour processing abil-

ities of infants at both 16 and 19 months by using pseudoisochromatic plates to force them to establish the contour around the shape. The results demonstrate that contour processing abilities may be sufficient at 16 months to allow participants to distinguish stimuli in that way. However, the much slower processing speed in contour processing between the 16 month age group and the 19 month age group indicates that 16 months might be a lower limit for this level of contour processing.

Data Availability

All data and analysis code required to analyse this dataset are available at https://osf.io/j8657/?view_only=64702cf578ce4ab89e9d884759281a87.

References

- Aisch, G. (2011). *Generating Color Blindness Test Images with Processing*. Retrieved from www.vis4.net
- Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. A. (2019). Rain-cloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Research*, 4, 1–41. doi: 10.12688/wellcomeopenres.15191.1
- Althaus, N., & Westermann, G. (2016). Labels constructively shape object categories in 10-month-old infants. *Journal of Experimental Child Psychology*, 151, 5–17. doi: 10.1016/j.jecp.2015.11.013
- Atkinson, J. (2002). *The Developing Visual Brain*. Oxford University Press. doi: 10.1093/acprof:oso/9780198525998.001.0001
- Atkinson, J., & Braddick, O. (1998). Research methods in infant vision. In *Vision research: A practical approach* (pp. 2–8). Oxford University Press.
- Atkinson, J., Braddick, O., & French, J. (1979). Contrast sensitivity of the human neonate measured by the visual evoked potential. *Investigative Ophthalmology & Visual Science*, 18, 210–213.
- Atkinson, J., Braddick, O., & Moar, K. (1977). Infants' detection of image defocus. *Vision Research*, 17, 1125–1126.
- Bergelson, E., & Aslin, R. N. (2017). Nature and origins of the lexicon in 6-month-olds. *Proceedings of the National Academy of Sciences*(Early Edition), 1–6. doi: 10.1073/pnas.1712966114
- Bornstein, M. H. (1985). On the development of color naming in young children: Data and theory. *Brain and Language*, 26(1), 72–93. doi: 10.1016/0093-934X(85)90029-X
- Bornstein, M. H., Kessen, W., & Weiskopf, S. (1976). The categories of hue in

- infancy. *Science*, 191(4223), 201–202. doi: 10.1126/science.1246610
- Braddick, O., & Atkinson, J. (2011). Development of human visual function. *Vision Research*, 51(13), 1588–1609. doi: 10.1016/j.visres.2011.02.018
- Brown, A. M., & Lindsey, D. T. (2013). Infant color vision and color preferences: A tribute to Davida Teller. *Visual Neuroscience*, 30(5-6), 243–250. doi: DOI: 10.1017/S0952523813000114
- Burkhalter, A., Bernardo, K., & Charles, V. (1993, 5). Development of local circuits in human visual cortex. *The Journal of Neuroscience*, 13(5), 1916–1931. doi: 10.1523/JNEUROSCI.13-05-01916.1993
- Bürkner, P.-C. (2017). brms : An R Package for Bayesian Generalized Linear Mixed Models using Stan. *Journal of Statistical Software*, 80(1), 1–28. doi: doi:10.18637/jss.v080.i01
- Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, 10(1), 395. doi: 10.32614/RJ-2018-017
- Chen, Y.-C., & Westermann, G. (2018). Different novelties revealed by infants' pupillary responses. *Scientific Reports*, 8(1), 1–8. doi: 10.1038/s41598-018-27736-z
- Dink, J. W., & Ferguson, B. (2018). *eyetrackingR*. Retrieved from <http://www.eyetrackingr.com>
- Fernald, A., Pinto, J. P., Swingle, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, 9(3), 228–231. doi: 10.1111/1467-9280.00044
- Fernald, A., Thorpe, K., & Marchman, V. A. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective-noun phrases. *Cognitive Psychology*, 60(3), 190–217. doi: 10.1016/j.cogpsych.2009.12.002
- Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while lis-

- tening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernandez, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). Amsterdam/Philadelphia: John Benjamins Publishing Company. doi: 10.1075/lald.44.06fer
- Forbes, S. H., & Plunkett, K. (2019a). Infants show early comprehension of basic color words. *Developmental Psychology*, 55(2), 240–249. doi: 10.1037/dev0000609
- Forbes, S. H., & Plunkett, K. (2019b). The role of colour labels in mediating toddler visual attention. *Cognition*, 186, 159–170. doi: 10.1016/j.cognition.2019.01.008
- Forbes, S. H., & Plunkett, K. (2020). Linguistic and cultural variation in early color word learning. *Child Development*, 91(1), 28–42. doi: 10.1111/cdev.13164
- Franklin, A., & Davies, I. R. L. (2004). New evidence for infant colour categories. *British Journal of Developmental Psychology*, 22(3), 349–377. doi: 10.1348/0261510041552738
- Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008, 12). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, 2(4), 1360–1383. doi: 10.1214/08-AOAS191
- Gerhardstein, P., Kovacs, I., Ditre, J., & Feher, A. (2004, 12). Detection of contour continuity and closure in three-month-olds. *Vision Research*, 44(26), 2981–2988. doi: 10.1016/j.visres.2004.06.023
- Gerhardstein, P., Tse, J., Dickerson, K., Hipp, D., & Moser, A. (2012, 10). The human visual system uses a global closure mechanism. *Vision Research*, 71, 18–27. doi: 10.1016/j.visres.2012.08.011
- Ghosh, J., Li, Y., & Mitra, R. (2018). On the Use of Cauchy Prior Distributions

- for Bayesian Logistic Regression. *Bayesian Analysis*, 13(2), 359–383. doi: 10.1214/17-BA1051
- Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, 14(01), 23. doi: 10.1017/S030500090001271X
- Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language*, 27(3), 689–705. doi: 10.1017/S0305000900004414
- Harris, L., Atkinson, J., & Braddick, O. (1976). Visual contrast sensitivity of a 6-month-old infant measured by the evoked potential. *Nature*, 264, 570–571.
- Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, 13(2), 128–157. doi: 10.1080/15250(xH)701795598
- Ishihara, S. (1917). *Tests for Color-Blindness*. Tokyo: Hongo Harukicho.
- Kanizsa, G. (1987). Quasi-Perceptual Margins in Homogeneously Stimulated Fields. In S. Petry & G. E. Meyer (Eds.), *The perception of illusory contours* (pp. 40–49). New York, NY: Springer New York. doi: 10.1007/978-1-4612-4760-9{_}4
- Kay, M. (2020). *tidybayes: Tidy Data and Geoms for Bayesian Models*. Retrieved from <http://mjskay.github.io/tidybayes/> doi: 10.5281/zenodo.1308151
- Knoblauch, K., Vital-Durand, F., & Barbur, J. L. (2001). Variation of chromatic sensitivity across the life span. *Vision Research*, 41(1), 23–36. doi: 10.1016/S0042-6989(00)00205-4
- Meints, K., Plunkett, K., & Harris, P. L. (1999). When does an ostrich become a bird? The role of typicality in early word comprehension. *Developmental Psychology*, 35(4), 1072–1078. doi: 10.1037/0012-1649.35.4.1072

- Mercer, M. E., Drodge, S. C., Courage, M. L., & Adams, R. J. (2014). A pseudoisochromatic test of color vision for human infants. *Vision Research*, 100, 72–77. doi: 10.1016/j.visres.2014.04.006
- Mirman, D. (2014). *Growth Curve Analysis and Visualization Using R Analysis and Visualization Using R*. Boca Raton, FL: Chapman & Hall / CRC Press.
- Patel, A., Maurer, D., & Lewis, T. L. (2010). The development of spatial frequency discrimination. *Journal of Vision*, 10(14), 1–10. doi: 10.1167/10.14.1
- Pease, P. L., & Allen, J. (1988). A new test for screening color vision: Concurrent validity and utility. *American Journal of Optometry & Physiological Optics*, 65(9), 729–738.
- Peeples, D. R., & Teller, D. Y. (1970). Color vision and brightness discrimination in two-month-old human infants. *Science*, 189.
- Peeples, D. R., & Teller, D. Y. (1978). White-adapted photopic spectral sensitivity in human infants. *Vision Research*, 18, 49–53.
- Plunkett, K., Hu, J.-F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition*, 106, 665–681. doi: 10.1016/j.cognition.2007.04.003
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org/>
- Stan Development Team. (2016). *RStan: the R interface to Stan*.
- Swingle, D. (2010). Fast mapping and slow mapping in children's word learning. *Language Learning and Development*, 6, 179–183. doi: 10.1080/15475441.2010.484412
- Taylor, G., Hipp, D., Moser, A., Dickerson, K., & Gerhardstein, P. (2014). The development of contour processing: Evidence from physiology and psy-

- chophysics. *Frontiers in Psychology*, 5, 1–10. doi: 10.3389/fpsyg.2014.00719
- Teller, D. Y. (1998). Spatial and temporal aspects of infant color vision. *Vision Research*, 38(21), 3275–3282.
- Teller, D. Y., Civan, A., & Bronson-Castain, K. (2004). Infants' spontaneous color preferences are not due to adult-like brightness variations. *Visual Neuroscience*, 21(3), 397–401. doi: 10.1017/S0952523804213360
- Wagner, K., Dobkins, K., & Barner, D. (2013). Slow mapping: Color word learning as a gradual inductive process. *Cognition*, 127(3), 307–317. doi: 10.1016/j.cognition.2013.01.010
- Wagner, K., Jergens, J., & Barner, D. (2018). Partial color word comprehension precedes production. *Language Learning and Development*, 1–21. doi: 10.1080/15475441.2018.1445531
- Younger, B. A. (1985). The segregation of items into categories by ten-month-old infants. *Child Development*, 56(6), 1574–1583.
- Yurovsky, D., & Frank, M. C. (2017, 3). Beyond naïve cue combination: salience and social cues in early word learning. *Developmental Science*, 20(2), e12349. doi: 10.1111/desc.12349