THE OFFICIAL JOURNAL OF THE
INTERNATIONAL CONGRESS
OF INFANT STUDIES

WILEY Blackwell

Infancy, 1–24, 2018

ISSN: 1525-0008 print / 1532-7078 online
DOI: 10.1111/infa.12260

A Novel Method to Investigate How Dimensions Interact to Inform Perceptual Salience in Infancy

Marie Rogers, and Anna Franklin
The Sussex Colour Group
School of Psychology
University of Sussex

Kenneth Knoblauch
Univ Lyon
Université Claude Bernard Lyon 1
Inserm, Stem Cell and Brain Research Institute U1208

How physical dimensions govern children's perception, language acquisition, and cognition is an important question in developmental science. Here, we use the psychophysical technique of maximum likelihood conjoint measurement (MLCM) as a novel approach to investigate how infants combine information distributed along two or more dimensions. MLCM is based on a signal detection model of decision that allows testing of several models of how observers integrate information to make choices. We tested 6-month-old infants' preferential looking to "green" stimuli that covaried in lightness and chroma and analyzed infant preferences using MLCM. The findings show that infant looking is driven primarily by lightness, with darker stimuli having a greater preference than lighter, plus a small but significant positive contribution of chroma. This study demonstrates that the technique of MLCM can be used in conjunction with preferential looking to investigate the salience of physical dimensions during development. The technique could now be applied to investigate the role of physical dimensions in key aspects of perceptual and cognitive development such as face recognition, language acquisition, and object recognition.

Our world is multidimensional. Objects vary along multiple physical dimensions, for example, apples vary in their glossiness, color, shape, size, and texture. Dimensional properties of an object covary and interact to produce cohesive perceptual experiences. However, it can be difficult to identify how physical dimensions govern perception, cognition, and behavior. This is important to understand perceptual and cognitive development such as language acquisition, object individuation, esthetics, and memory.

Correspondence should be sent to Anna Franklin, The Sussex Colour Group, School of Psychology, University of Sussex, Brighton, BN1 9QH, UK. E-mail: anna.franklin@sussex.ac.uk

Perceptual salience of physical dimensions

During development, some dimensions are particularly salient in a given context. In language acquisition, children exhibit certain attentional biases that lead them to prioritize learning about one dimension over another, depending on the circumstance. For example, in an object label extension task, children display a "shape bias" (Landau, Smith, & Jones, 1988). When children learn a novel word, they are more likely to extend that word to similarly *shaped* objects, versus objects of a similar color (Bornstein, 1985), a similar texture or color (Diesendruck & Bloom, 2003), or a similar size or texture (Landau et al., 1988). However, there is also evidence that such attentional biases are task-dependent, for example, children are more likely to extend a novel word when an object shares a unique *function* than shape (Diesendruck, Markson, & Bloom, 2003). The most salient dimension to a child during novel word learning can be context-dependent. These examples demonstrate the importance of understanding dimensional salience in the domain of language and having an effective task to test this.

Dimensional properties also contribute to object recognition and individuation during development. In occlusion events, where an object is hidden behind a screen, infants are faced with the challenge of *object individuation* when the object reappears. They must determine how many objects are moving in and out of view. Objects have dimensional (or featural) properties, which may help solve this puzzle, such as shape, size, texture, and color. If the reappearing object has different properties to the original one, an adult would be expected to assume that this object is distinct from the occluded object. The dimensions which infants use to solve this puzzle change over the course of development: Infants at 4.5 months use shape and size features, at 7.5 months they additionally use pattern, and at 11.5 months they can use color in object individuation (Wilcox, 1999). Covariation of multiple dimensional properties further aids in object individuation. When the color *or* luminance of an object varies, infants are not able to individuate the object until 11.5 months (Wilcox, 1999), but with covariation, infants aged 7.5 months are able to complete this task (Woods & Wilcox, 2010). This may be because color and luminance often covary in our natural visual environment and thus are useful cues for identification.

Interaction of dimensions

The evidence discussed above pitted dimensions against each other and examined which dominated as the most important (salient) in a given task. In reality, it is possible that, when multiple physical dimensions covary, perception and behavior arise from an interaction between the dimensions. For example, in adults, the dimensions of color (hue, chroma, and lightness) interact in perception (Burns & Shepp, 1988; Rogers, Knoblauch, & Franklin, 2016). However, a study which modeled the data from five studies on infants' preferential looking at colors found that 6-month-old infants tended to base their looking preference on hue, and luminance differences did not contribute to the fit of the model (Brown & Lindsey, 2013). This raises the interesting question of whether there are developmental differences in how the dimensions of color are weighted in perceptual judgments.

It has also been argued that infants process stimuli more holistically than adults do and that there is a developmental trend from interaction toward separation of dimensions (e.g., Kemler, 1983; Kemler & Smith, 1978, 1979; Smith, 1979, 1983; Smith & Kemler, 1978). That is, dimensions, which adults separate (such as shape and color), are more integrated into infants' perception. However, research into this hypothesis has yielded mixed results. For example, while there is evidence to support this hypothesis (e.g., Kemler, 1983), one study provides contrary evidence and found that younger children (aged 5–6 years) did not use holistic rules on a dimensional card-sorting task to a greater extent than older children (aged 10–11 years; Kemler & Smith, 1978).

The most appropriate and effective method to investigate dimensional processing in infants is not clear. A method is needed which can quantify the relative contributions of multiple dimensions to perceptual behaviors (e.g., preferential looking) in infancy. Studies conducted with adults on the interaction of dimensions require explicit judgments, such as odd-one-out tasks or similarity ratings (e.g., Burns & Shepp, 1988; Indow & Kanazawa, 1960; Komarova & Jameson, 2013), but these explicit judgment methods are not feasible for infants. In infants, modeling preferential looking with regressions is useful in identifying which dimension contributes to preferential looking (e.g., Brown & Lindsey, 2013), but the method does not address the extent to which multidimensional interactions may occur.

Furthermore, when determining how two (or more) physical dimensions contribute to behavior, there is a challenge of equating the dimensions on a perceptual scale. For example, at 9 months, infants appear to identify objects based on shape but not color (e.g., Kaldy & Leslie, 2003), but how can we be sure that the shapes selected in the experiment were equally salient to the colors? This has been addressed using interdimensional salience mapping (ISM; Kaldy, Blaser, & Leslie, 2006; Kaldy & Blaser, 2009, 2013) which allows one to determine how physical change along a dimension relates to changes in salience and thus enables salience to be equated across dimensions. ISM uses a forced-choice preferential-looking method to determine which of two stimuli are more salient in a head-to-head competition. By manipulating the properties of the stimuli and determining which of the two competitors "wins" in such a task, ISM can produce a psychometric function of salience. This precise approach to stimulus calibration across dimensions is a big improvement on the majority of previous developmental studies of dimensional processing, which do not consider stimulus calibration issues. However, although ISM can equate perceptual salience across dimensions, we still lack a method that quantifies dimensional *interactions* in perception.

Maximum likelihood conjoint measurement

Here, we present a novel approach to investigate the interaction of color dimensions in infancy. MLCM is a psychophysical technique that allows one to quantify the contribution of more than one dimension to a behavior (Ho, Landy, & Maloney, 2008; Knoblauch & Maloney, 2012). MLCM determines how the probability of a choice between a pair of stimuli is influenced by the covariation along multiple dimensions or attributes of the stimuli. Previous adult studies using MLCM have examined the interactions of gloss and surface texture (Ho et al., 2008; Qi, Chantler, Siebert, & Dong, 2015), surface lightness and gloss (Hansmann-Roth & Mamassian, 2017), contour curvature and luminance in the illusory watercolor effect (Gerardin, Devinck, Dojat, & Knoblauch, 2014), the voice and the face in gender perception (Abbatecola, Gerardin, Knoblauch, & Kennedy, 2016), race on the perceived lightness of faces (Nichiporuk, Knoblauch, Abbatecola, & Shevell, 2017), and lightness and chroma in adults (Rogers et al., 2016). In each of these experiments, adult participants made explicit judgments about the stimuli. However, MLCM has not yet been used to measure the interaction between dimensions in infancy and throughout development. To implement this approach in infants, we will use preferential eye movements on presentation of a pair of stimuli as the choice response required to implement MLCM.

In MLCM, the choice probabilities are modeled by a noise-contaminated decision rule. This constitutes a signal detection model whose parameters are estimated via maximum likelihood, hence the name of the method. Three nested decision rules can be defined. First, the *independent model* describes the case where the choice probabilities can be described by physical manipulation of one dimension alone; there is no "contamination" by the other dimension. This would occur, for example, if given a choice between a chromatic and an achromatic stimulus, the observer always chose to look at the chromatic one, regardless of the difference in lightness between the two. The *additive model* describes the case where choice probabilities depend on the additive sum of two underlying, response components, one associated with each of the manipulated dimensions. In this case, the choice is influenced by both variables, but the contribution of each variable to the choice depends only on its own level in the stimulus and not at all on the level of the other variable. Finally, in the *saturated model*, the choice probabilities depend on an interaction beyond the additive combination of the underlying components. Thus, like an interaction in an analysis of variance, the probabilities depend on the particular pair of values along each dimension rather than the sum of two components. The model is called saturated because it includes the maximum number of parameters to model the data. In each case, the decision variable is assumed to be perturbed on a trial-by-trial basis by mean zero, Gaussian noise. This allows for a stochastic relation between the responses and the decision variable. The noise-contaminated decision variable is related to the response through a Gaussian psychometric function. Importantly, relative judgments across the dimensions are equated through their effects on the decision variable, that is, the scales are constructed so that chroma and lightness differences that produce the same difference on the decision variable will be equally salient.

Current study

In this study, we use MLCM to estimate the influence of lightness and chroma on infants' preferential looking to color. The study aims to establish whether it is possible to use MLCM to model the contribution of physical dimensions to infants' preferential looking using the technique. We use color as a testing ground for the application of MLCM to infant preferential looking and to further our understanding of infants' color perception. By at least 3 months of age, infants are trichromatic (the three types of cone photoreceptors and the "red-green" and "blue-yellow" neural pathways are functioning; Banks & Bennett, 1988; Knoblauch, Bieber, & Werner, 1998; Morrone, Burr, & Fiorentini, 1990, 1993; Volbrecht & Werner, 1987). The ability to discriminate colors progressively improves through development until adolescence (Knoblauch, Vital-Durand, & Barbur, 2001). Previous studies have shown that infants look longer at some colors than others (e.g., Franklin, Bevis, Ling, & Hurlbert, 2010; Franklin et al., 2008; Zemach, Chang, & Teller, 2007) and can categorize the spectrum of color (e.g., Skelton, Catchpole, Abbott, & Franklin, 2017). However, we know little about how the dimensions of color contribute and interact in infants' perception of color.

The appearance of color can be described in a three-dimensional perceptual space consisting of *hue*, *lightness*, and *chroma* (Wyszecki & Stiles, 2000). As discussed, these dimensions are not independent in adults' perception (Burns & Shepp, 1988; Rogers et al., 2016). It is unknown whether this is the case for infants as well, or whether there are developmental differences in the interaction of color dimensions.

To test whether MLCM can be applied to infants' preferential looking and to further understand infant color perception, we conducted an experiment using a forced-choice preferential-looking method with 6-month-olds and adult observers (Teller, 1979). The method involves eye-tracking observers' responses to pairs of stimuli and coding which of the pair they look at first. For infants, the pairs of stimuli were randomly selected from a 3-by-3 matrix of green stimuli in which lightness and chroma independently covary. All stimulus levels were above threshold for infants at 6 months according to threshold data from Knoblauch et al. (2001). The adult experiment used the same range of lightness and chroma, but the range was divided into four stimulus levels rather than three. We were able to use more stimulus levels with adults as they have better color discrimination and can tolerate a longer experimental procedure.

We analyzed the data using MLCM to estimate the relative contributions of lightness and chroma to the observers' decisions (first look). The outcome variable of our MLCM analysis is parameterized to be on the scale of *d'* (units of the standard deviation for each scale value) from signal detection theory (e.g., Gerardin et al., 2014; Green & Swets, 1966; Ho et al., 2008; Rogers et al., 2016; Stanislaw & Todorov, 1999). Here, we take *d'* to reflect the perceptual salience of a stimulus, not necessarily its discriminability, by measuring which of two stimuli observers look at first. We follow Kaldy and Blaser's definition of salience as, "the visual system's real-time assessment of the behavioural relevance (current importance) of information in the scene—a prioritization that drives attention allocation and consequent eye movements" (2009, p. 223). The most salient object in a scene is the one that is preferred, that is, it beats the other in a forced-choice looking paradigm. Note that what we aim to measure is the suprathreshold salience and contributions of chroma and lightness when these are dimensions of a multifeature object, not the low-level detectability of chroma and lightness as separate dimensions. Although it has been argued that detectability determines visual salience in infancy (e.g., Banks & Salapatek, 1981), the lack of transitivity in infant visual salience suggests that infants' visual salience is not purely driven by detectability of each dimension (Kibbe, Kàldy, & Blaser, 2018). With MLCM, we can quantify multidimensional contributions to perceptual salience.

METHOD

Participants

Twenty-two 6-month-olds (nine females) participated in this study in total. One further infant was recruited but did not take part due to fussiness. The infants had a mean age of 28 weeks (SD = 1.9) and all had a birthweight >2.5 kg, with no family history of color deficiency, and no known visual impairments. Infant participants were recruited by contacting parents/carers with infant children through the Sussex Baby Lab (University of Sussex, UK). They received a small gift (book or T-shirt) as a thank you at the end of the experiment.

Additionally, 12 adult observers participated (all female) with a mean age of 22 years (*SD* = 3.5). All adult observers were assessed as having normal color vision using Ishihara plates (Ishihara, 2010) and the Lanthony Tritan Album (Lanthony, 1998). Adult participants were paid £8 per hour for their participation. This study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any data collection. All procedures involving human subjects in this study were approved by the Sciences and Technology Cross-Schools Ethical Committee at the University of Sussex and the European Research Council Executive Agency ethics committee.

Stimuli and apparatus

Three levels of lightness and three levels of chroma were selected for the infant stimuli, giving a 3-by-3 stimulus matrix (see Table 1 and Figure 1). The levels were specified in CIE LCH_{uv} color space. This is a transformed version of CIELUV space, where L is the lightness, C is the chroma, and H is the hue (Poynton, 2012). This color space was selected to be in line with a previous study of MLCM in adults (Rogers et al., 2016). The hue angle was fixed at 143.2° (CIE H), which normal adult observers classify as "green."

We maximized the stimulus range within monitor gamut (i.e., the possible range of colors that can be displayed on the screen) to maximize discriminability for infants. Equal perceptual spacing in CIE LCH_{uv} space between adjacent stimuli was calculated from previously obtained results from adults by Maximum Likelihood Difference Scaling (MLDS) (Rogers et al., 2016). MLDS is a psychophysical scaling technique, also based on a signal detection model of decision, that allows the estimation of an interval scale along a continuous dimension through comparisons of stimulus intervals (Knoblauch & Maloney, 2008; Maloney & Yang, 2003). The derived scale has the property that equal scale differences appear equally different perceptually. The prescaling of the stimuli with MLDS is not a prerequisite to performing or to analyzing the MLCM results. It serves mostly as a convenience to linearize the estimated response functions, which simplifies the subsequent statistical analyses.

The adult stimuli used the same range of values, but divided the range into four levels, giving a 4-by-4 stimulus matrix. A gray background (xyY (1931): 0.313 0.329, 50; L^* = 100) was used throughout the experiment, lighter than all stimulus levels.

Stimuli were presented on a 22-inch Mitsubishi Diamond Plus 230SB monitor, calibrated using a ColorCAL colorimeter (Cambridge Research Systems), and the room was dark other than that light emitted from the monitor. Eye movements were

TABLE 1
Lightness and Chroma Values for Adult and Infant Stimuli in CIE LCH_{uv}. Hue Angle = 143.2°

	Infant stimuli			*Adult stimuli*	
Stimulus level	*Chroma*	*Lightness*	*Stimulus level*	*Chroma*	*Lightness*
1	5.00	39.63	1	5.00	39.63
2	26.83	52.55	2	19.11	49.55
3	50.00	69.06	3	34.85	58.47
			4	50.00	69.06

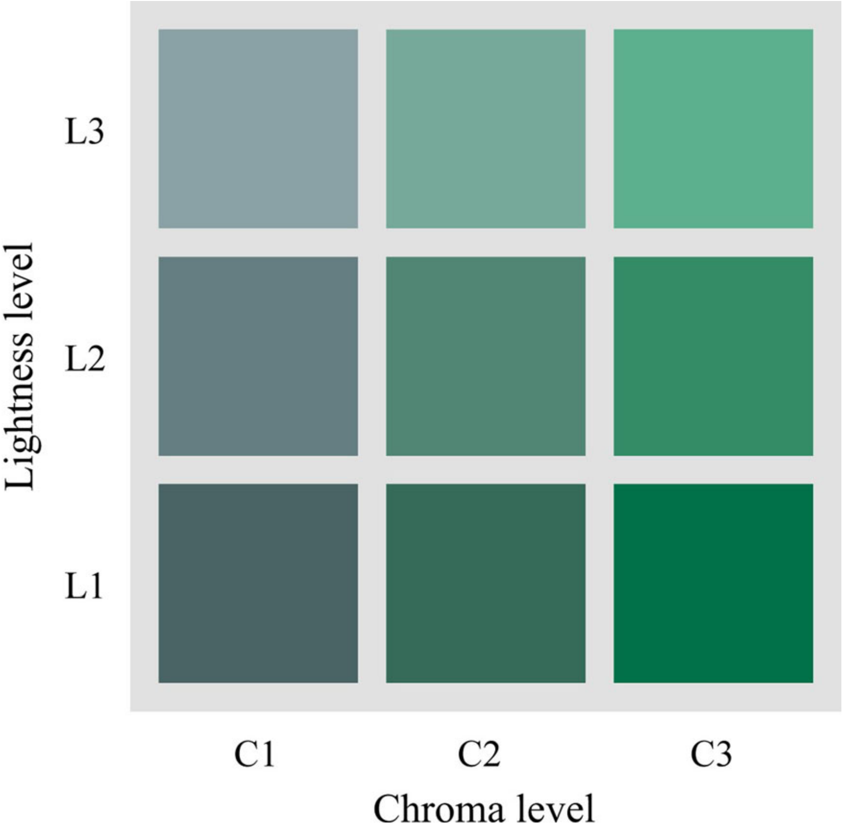

Figure 1 The 3-by-3 matrix of infant stimuli. Three levels of chroma and three levels of lightness are perceptually equally spaced using MLDS.

recorded with an EyeLink 1000 Plus Eye Tracker manufactured by SR Research Ltd, using an infant lens. Participants wore a small target sticker, which aids accurate tracking with a freestanding eye tracker. The experimental procedure was created using SR Research Experiment Builder.

Design and procedure

On each trial, a pair of stimuli was chosen from the 3-by-3 stimulus matrix (see Figure 1). All possible unordered pairs were shown in a randomized order, excluding self-comparisons. From the nine stimuli, there are 72 pair combinations, including both left/right and right/left versions of a pair ($9 \times (9 - 1) = 72$).

Infants sat in a car seat 60 cm away from the computer monitor. Cartoon clips played onscreen while the researcher set up the eye-tracking camera; infants then completed a 3-point calibration. On each trial, a pair of stimuli was presented: one on the left and one on the right-hand side of the screen. Each stimulus was 14 cm × 14 cm on the monitor, corresponding to a visual angle of 13°. If the infant did not fixate on either the left or the right-hand stimulus before the trial timed out (2,000 ms), the stimulus pair was presented again later in the experiment. Thus, there were, in

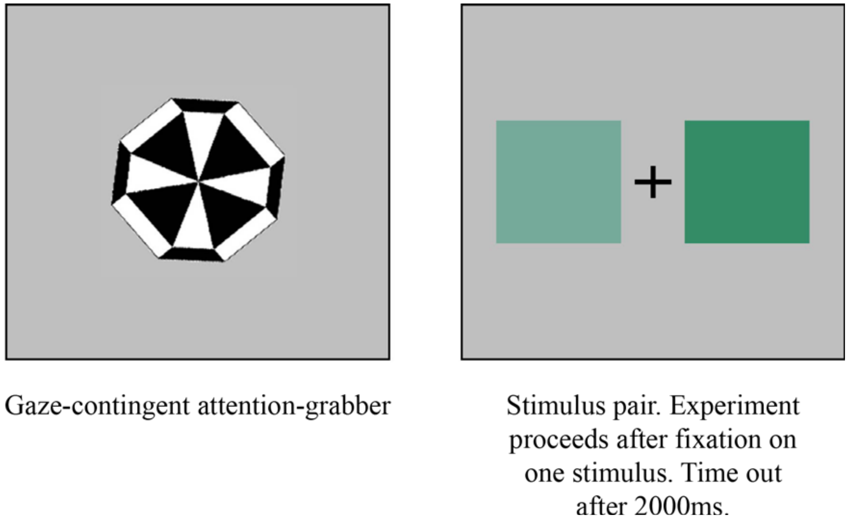

Figure 2 Trial procedure.

principle, 72 ordered pairs of colors presented, but the number of total trials for an individual observer could be higher as a pair of stimuli would be reshown until the infant fixated on one side.

To begin, a looming black and white "attention-grabber" appeared with a noise, to ensure the infants were centrally fixated before the trial began. The trial began automatically when the participant fixated on the attention-grabber. The experiment automatically proceeded to the next screen when the infant fixated on either the left or the right stimulus for 160 ms (this fixation duration has been found to be reliable and resistant to measurement artifacts; Wass, Smith, & Johnson, 2013). See Figure 2 for illustration of trial procedure. The experiment was halted if the infant showed signs of distress. Infants completed an average of 68% of the total number of trials (SD = 29%).

Adult participants were informed that they were a comparison group for an infant experiment and to "look at the patch that stands out more, or that most grabs your attention." They completed a 9-point calibration before the trial began. The procedure was the same as the infant experiment, except there was a central fixation cross instead of the attention-grabber. Furthermore, as adults viewed unordered pairs randomly selected from a 4-by-4 stimulus matrix (i.e., 16 stimuli), they responded to 240 trials ($16 \times (16 - 1) = 240$).

Data analysis

For each participant and on each trial, the stimulus looked at first was recorded as the choice. The data were analyzed using MLCM; this analysis is described in full elsewhere (Knoblauch & Maloney, 2012; Rogers et al., 2016), and we provide a brief summary here. On each trial, a pair of stimuli is randomly selected from the stimulus matrix (Figure 1) for presentation. The pair can be indexed in terms of their ordinal chroma levels (w, y) and lightness levels (x, z). For example, in the pair C3L1 (bottom right stimulus) plus C2L2 (center stimulus), $w = 3$, $y = 2$, $x = 1$, and $z = 2$. It is assumed that when viewing the stimulus pair, the observer forms the noise-contaminated decision variable:

$$\Delta = \delta + \varepsilon = \Psi\left(\varphi_y^C, \varphi_z^L\right) - \Psi\left(\varphi_w^C, \varphi_x^L\right) + \varepsilon, \tag{1}$$

where $\delta = \Psi\left(\varphi_y^C, \varphi_z^L\right) - \Psi\left(\varphi_w^C, \varphi_x^L\right)$ and the Ψ terms are internal responses that depend on the contributions of the stimulus lightness and chroma to the perceptual salience. The value of Δ predicts the first-look response, that is, the observer looks left if it is greater than 0 and right if it is less. The noise term, ε, assumed to be zero-mean Gaussian with variance σ^2, is included to account for the fact that observers do not necessarily make the same choice on repeated trials. MLCM is an equal variance, Gaussian, signal detection model and allows estimation of the scale values corresponding to the contributions of each internal response by maximum likelihood so that they best predict the observer's behavior over the set of experimental responses.

As described above, there are three possible nested models of the decision variable that can be fit to the data: *independent*, *additive,* and *saturated*. With the independent model, the observer's responses depend on only one of the dimensions. The decision rule reduces to

$$\Delta(w, x, y, z) = \Psi_y^C - \Psi_w^C + \varepsilon \tag{2}$$

in the case of a chroma response or

$$\Delta(w, x, y, z) = \Psi_z^L - \Psi_x^L + \varepsilon \tag{3}$$

in the case of lightness. With k levels of each dimension, there are $k + 1$ parameters including the variance for the noise term. However, to make the model identifiable, the lowest level is set to 0 and the variance to 1, so that there are only $k - 1$ free parameters to estimate.

For the additive model, the decision variable becomes

$$\Delta(w, x, y, z) = \left(\Psi_y^C + \Psi_z^L\right) - \left(\Psi_w^C + \Psi_x^L\right) + \varepsilon. \tag{4}$$

With k levels along each dimension, there are $2k + 1$ parameters, including the variance for the noise term. To make the model identifiable, the two lowest levels along each dimension are set to 0 and the variance to 1, yielding $2k - 2$ free parameters to estimate. Finally, for the saturated model, the decision variable becomes

$$\Delta(w, x, y, z) = \left(\Psi_y^C + \Psi_z^L + \Psi_{yz}^{CL}\right) - \left(\Psi_w^C + \Psi_x^L + \Psi_{wx}^{CL}\right) + \varepsilon. \tag{5}$$

Due to the interaction terms, the responses cannot be explained by a simple additive sum of component responses, but depend on the specific levels along each dimension. With k levels along each dimension, there are $k^2 + 1$ parameters, including the variance for the noise term. In order to make the model identifiable, only one cell in the $k \times k$ grid is set to zero and the variance to one, yielding a model with $k^2 - 1$ free parameters to estimate. This is the maximum number, hence the term saturated.

If we denote the chroma and lightness quadruple of indices (w, x, y, z) for a trial by q, then assigning responses, R, the values 0/1 to choices left/right, respectively, the probability of choosing the right-hand stimulus on a trial can be written as

$$P(R = 1) = \Phi\left(\frac{\delta_q}{2}\right)^{R_q}\left(1 - \Phi\left(\frac{\delta_q}{2}\right)\right)^{1-R_q} \tag{6}$$

where Φ is the cumulative distribution function of the standard Gaussian, the value of 2 in the denominator assures a unit variance for each value of Ψ, and the log likelihood of the set of responses over all trials is

$$\ell(\Psi; R, q) = \sum_q \log(P_q). \tag{7}$$

We choose the values of Ψ to maximize the likelihood over the set of responses of the observer by minimizing the negative of the expression above for each of the three models. In practice, the model can be reformulated as a generalized linear model (GLM) with a binomial family (McCullagh & Nelder, 1989),

$$g(E[R = 1]) = X\beta, \tag{8}$$

where the link function, g, is the inverse cumulative Gaussian distribution function, X, is a design matrix with a column for each identifiable term in the model filled with 0, -1, or 1 depending on the stimulus levels on the trial and the sign of the coefficients in the decision rule and β is a vector of the estimated identifiable scale values. The three nested models are tested with likelihood ratio tests. The statistic,

$$-2(\ell_O - \ell_1) = \chi^2_{df}, \tag{9}$$

where ℓ_0 and ℓ_1 are maximum likelihoods from nested models, is distributed asymptotically according to a chi-square distribution with degrees of freedom the difference of number of parameters between the two models (Wood, 2015). We fit the models to the data and tested them using the MLCM package in the open-source software R (Knoblauch & Maloney, 2014; R Core Team, 2017) that implements the above GLM and testing procedures.

RESULTS

Infants

Six of the 22 infants completed all 72 trials, and 11 completed over 75% of the trials (range of trials completed = 13–72). The analysis of one infant observer (number 8) showed complete separation. This occurs where the outcome variable is perfectly predicted by an explanatory variable (Albert & Anderson, 1984). Further investigation showed that in trials where the highest chromatic value was presented, the infant fixated on this stimulus 100% of the time. While this appears to be prima facie evidence that the infant responded to chromatic stimuli, it presents a problem for analysis because we cannot obtain a valid estimate of the dimension's contribution without variation in response. Therefore, we (conservatively) excluded this infant's data from the full analysis. Analyses reported here are based on data from 21 infants.

The infants' frequencies of first-look responses to the green stimuli are plotted in Figure 3. The figure shows that within each chroma level, darker stimuli tend to be more preferred on the first look than lighter stimuli. There also appears to be an effect of chroma, with higher chroma levels slightly more preferred, especially at the highest and lowest lightness levels. A log-linear analysis of the frequencies shows the chroma:-lightness interaction fails to reach significance ($\chi^2(4) = 0.86, p = .93$) nor does the main effect of chroma ($\chi^2(2) = 0.09$). Nevertheless, the model that included main effects of lightness and chroma displayed a lower AIC than one without the chroma term (AIC (L + C) = 70.8; AIC(L) = 71.5) suggesting that the model with the chroma term is closer to the true model. These results were obtained by combining responses across all infants and ignoring observer dependent variations, thus, reducing the power of the analysis. We address this issue more directly in the next set of analyses.

We used MLCM to fit the additive model to the responses of each individual infant to estimate the contributions of each dimension to the looking preferences. The points in Figure 4 show the average additive contributions of lightness and chroma to infants' looking preference and the 95% confidence intervals. The y-axis indicates d', which is a measure of sensitivity or response strength used in signal detection theory, and that here indicates perceptual salience. Figure 4 shows that lightness negatively

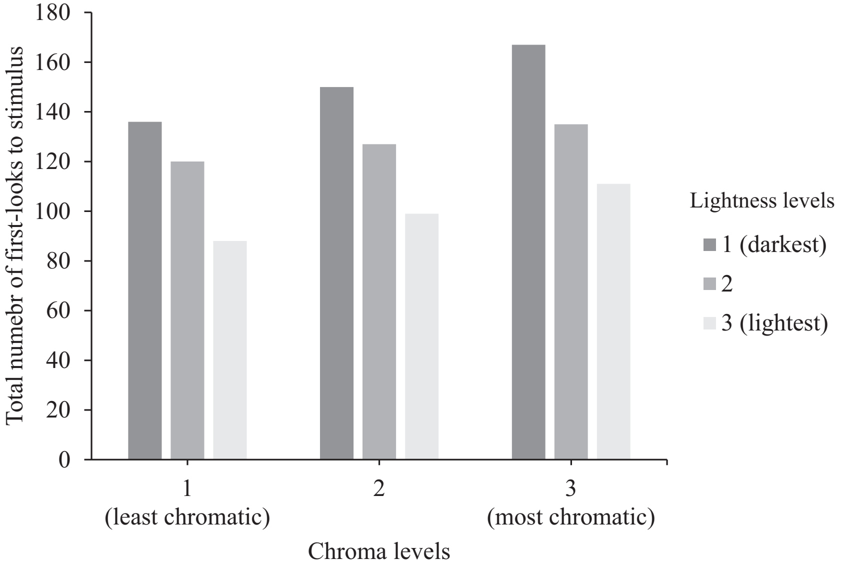

Figure 3 Frequency of first-look responses to each of the stimuli in the 3-by-3 stimulus matrix, summed across all infant observers. N = 21.

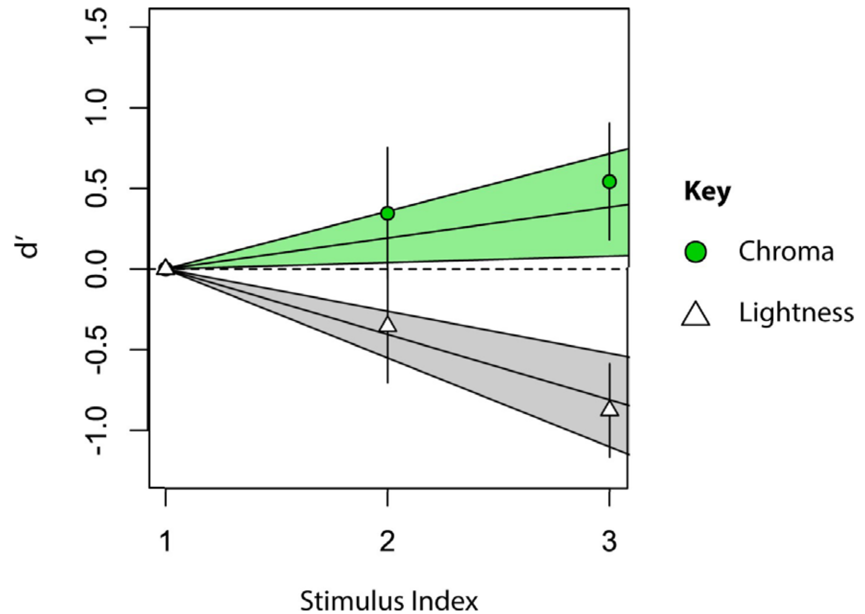

Figure 4 Parameter estimates for the additive model for looking preference in infants, averaged across observers. Circles show the estimated contribution of chroma and triangles of lightness based on individual MLCM fits. Error bars show 95% confidence intervals of the points. The lines are based on the mixed-effects model fixed-effect slope estimates, and the shaded areas, the 95% confidence intervals for the slopes of the lines. N = 21.

contributes to infants' first-look preference, that is, darker stimuli tend to be preferred over lighter stimuli. Conversely, chroma shows a smaller positive contribution to looking preference, indicating that stimuli with higher chromaticity are preferred over stimuli with lower chromaticity. The responses increase approximately linearly with the stimulus indices, suggesting that the stimulus spacing based on the adult scaling functions also holds approximately for the infants. While the estimated values of d' are small, the confidence intervals for the strongest lightness and chroma values indicate significant differences from zero.

A criticism that can be raised is that the number of trials to estimate the response probabilities is quite small. If all 72 trials were completed, then there are only two trials for each of the 36 unique pairs. As indicated earlier, not all infants completed 72 trials and one infant completed as few as 13, surely insufficient to estimate all of the choice probabilities. For comparison, previous studies in adults that used more levels for each dimension tested on the order of 1,000 or more trials. Knoblauch and Maloney (2012) have shown with simulation that the precision of estimates in MLCM is related to the square root of the number of trials tested. Comparing, for example, with the Gerardin et al. (2014) study that tested 1,500 trials per condition, we would see a reduction in precision of $\sqrt{72/1,500} = 0.22$, at best, compared to their results.

Mixed-effects models provide a possible solution to this problem in that observers are assumed to be sampled from a common population that shares common characteristics (Knoblauch & Maloney, 2012; Moscatelli, Mezzetti, & Lacquaniti, 2012). Information is pooled over all observers, weighted according to the information available from each observer's data, to obtain an optimal population estimate. This leads to shrinkage of the predictions of extreme observer's means toward the population mean, as the individual estimates are considered to borrow strength from each other.

Generalized linear mixed-effects models (GLMM) are GLMs in which the linear predictors are composed of fixed and random effect terms. Estimates are made of the fixed-effect coefficients and the variances of the random terms (Bates, Maechler, Bolker, & Walker, 2014). We can use the approximate linearity of the response estimates to specify a GLMM for MLCM (Rogers et al., 2016) that combines the data from all infants in an optimal fashion. We are assuming that this equal spacing based on difference scaling is also applicable to infants, and therefore, the assumption of linearity in response is valid. In short, for each stimulus pair, we compute the difference in levels within each dimension, *dC* and *dL*, and use these variables as covariates. The GLMM can then be expressed as

$$g(E(R=1)) = (\beta_C + b_{C,o})dC + (\beta_L + b_{L,o})dL + \epsilon, \tag{10}$$

where g is the probit (inverse cumulative Gausian) link function, β_C and β_L, fixed-effect slopes for chromatic and luminance contributions, $b_{C,o}$ and $b_{L,o}$, observer-specific random variations of the slope each assumed to be Gaussian random variables with $\mu = 0$ and variances σ_C^2 and σ_L^2, respectively, and ϵ is a standard Gaussian variable with $\mu = 0$ and $\sigma^2 = 1$. The mixed-effects models were fit using the *glmer* function in the lme4 package (Bates et al., 2014) within the software R (R Core Team, 2017). We can use the same framework to fit the independence model by dropping one of the terms, covariates, *dC* or *dL*, and a saturated model by including terms that are the product of the two covariates. The three models are then evaluated using nested likelihood ratio tests.

Likelihood ratio tests indicated that the additive model fit the data significantly better than the independence models for lightness ($\chi^2(1) = 19.79$, $p \ll .001$) and for chroma ($\chi^2(1) = 5.90$, $p = .015$). However, the saturated model was not an improvement over the additive model ($\chi^2(1) = 0.289$, $p = .591$). This suggests that both chroma and lightness contribute to infants' looking preference with no interaction between the two dimensions. The additive model revealed a significant negative effect of lightness ($\beta_L = -0.202$, $z = -5.76$, $p \ll .001$) and a smaller but still significant positive effect of chroma ($\beta_C = 0.096$, $z = 2.48$, $p = .013$) on looking preference. The 95% confidence intervals for the slopes, based on profile likelihoods, are chroma: (0.020, 0.179) and lightness: (−0.275, −0.130), both excluding zero. The variance associated with the chroma term was 2.6 times larger than that for the lightness term, or in other terms, it accounted for 72% of the interobserver variability. The lines drawn through the data points in Figure 4 are based on the estimated slopes from the additive model and appear to describe the average data well, although the points for the chromatic averages suggest some bias in the estimates, perhaps arising from the small number of samples for the individual estimates. This also supports the use of the equal perceptual spacing of stimulus levels based on adult MLDS data. The envelopes about each curve display the 95% confidence limits on the fitted lines. Again, while the predicted effects are small, they do support significant contributions of both dimensions to the infants' judgments with a contribution of lightness to performance roughly twice that of the contribution of chroma.

Adults

As with the infant observers, we plotted the combined adults' first-look response as a bar chart in Figure 5. It appears that adults are more likely to look at a stimulus as it becomes more chromatic. There also seems to be a positive, rather than a negative, influence of lightness, with lighter stimuli being more preferred, especially at the highest lightness level (4) compared to the other three. The log-linear analysis did not support a significant lightness:chroma interaction ($\chi^2(9) = 9.1, p = .43$). Both the lightness and chroma main effects were significant, however (lightness: $\chi^2(3) = 59.3, p \ll .001$; chroma: ($\chi^2(3) = 301.7, p \ll .001$).

Individual adult responses were analyzed using MLCM with the additive model, and the average values for the contributions of lightness and chroma components with 95% confidence intervals are plotted in Figure 6. Chroma has a large positive effect on looking preference: Adults are more likely to look first toward a stimuli with a high chroma level. Lightness also has a positive contribution to first-look preference in adults, although it appears smaller, and is perhaps significant only at the highest lightness level.

The GLMM described above to test the infant data was fit to the adult responses. Likelihood ratio tests rejected the independence model for lightness ($\chi^2(1) = 4.63$, $p = .031$) and for chroma ($\chi^2(1) = 10.94$, $p < .001$). However, the saturated model fit could not be differentiated statistically from the additive model fit ($\chi^2(1) = 0.017$, $p = .897$). This suggests that both chroma and lightness contribute additively to adults' looking preference without the need of an interaction term. The mixed-effects analysis revealed a significant positive effect of chroma ($\beta_C = 0.599$, $z = 4.19$, $p \ll .001$) and a smaller but still significant positive effect of lightness ($\beta_L = 0.226$, $z = 2.37$, $p = .018$)

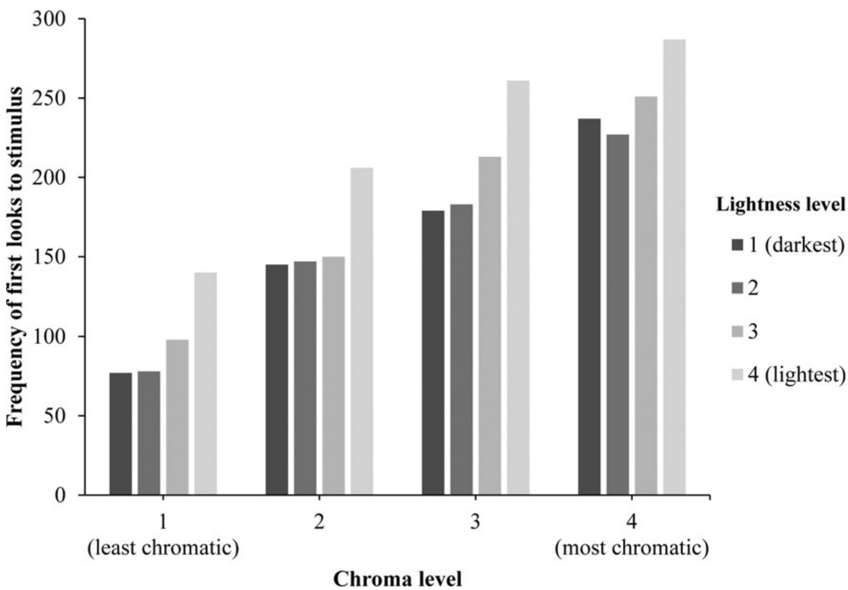

Figure 5 Frequency of first-look responses to each of the green stimuli in the 4-by-4 stimulus matrix, summed across adult observers. N = 12.

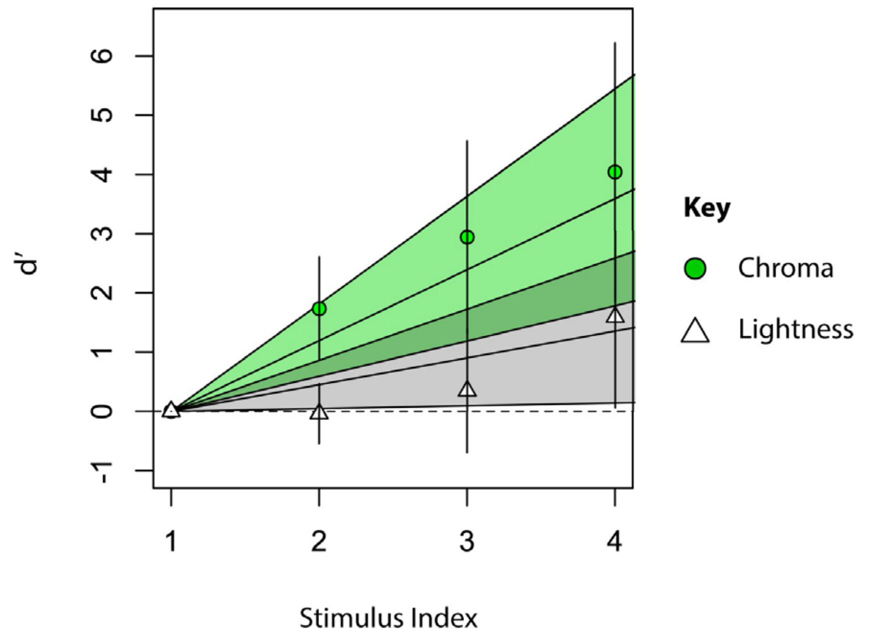

Figure 6 Parameter estimates for the additive model for looking preference in adults, averaged across observers. Circles show the estimated contribution of chroma and triangles of lightness. Error bars show 95% confidence intervals. N = 12.

on looking preference with 95% confidence intervals: chroma: (0.298, 0.908) and lightness: (0.024, 0.432). The lines drawn through the data points in Figure 6 are based on the estimated fixed-effect slopes from the additive model of the mixed-effects analysis. The variance associated with the chromatic term was 2.3 times larger than that for the lightness term, or in other terms, it accounted for 70% of the interobserver variability, similar to the infant value. The magnitude of the average effects is about three times larger than those of the infants. In adults, the contribution of chroma to performance is nearly three times that of lightness, that is, there is a reversal in the dominant dimension between adult and infant.

Influence of sampling constraints in infant data

The adults were tested with 16 stimuli generating 240 ordered pairs, while the infants were tested with only nine stimuli generating 72 ordered pairs. In addition, every adult completed all 240 trials while only about one quarter of the infants completed the full set of trials. The adult and infant data sets differ both quantitatively and qualitatively. To what extent could the differences in sampling account for these differences?

To address this question, we repeatedly simulated reduced and variable size data sets by sampling from the 12 adult data sets, to mimic the sampling characteristics of the infant data sets. We analyzed each simulated data set with a GLMM to assess whether the sampling influenced the magnitude and the slopes of the estimated chroma and lightness functions. First, we reduced the adult data sets to 72 trials using only nine stimuli with three levels per dimension by eliminating all trials that contained the fourth (highest) levels of chroma and/or lightness in the stimulus set. We will refer to

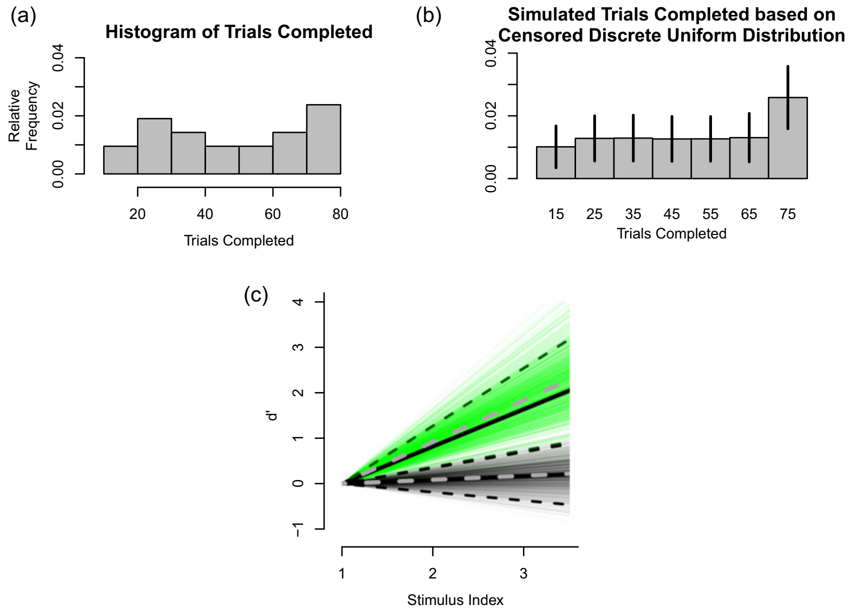

Figure 7 (a) Relative frequency distribution of number of trials completed for the 21 infants tested. (b) Average relative frequency distribution of number of trials completed based on a censored discrete uniform distribution with endpoints [13, 92] and with all values greater than or equal to 72 assigned to this value. One thousand samples of size 21 were drawn from the distribution. Each sample was assigned a bin according to the breaks in the abscissa of Figure 7a, and the averages were taken for each bin. Error bars are ±SD. (c) The results of the GLMM fit to the simulated data sets are shown as 1,000 thin green and black transparent lines with the slopes of the estimated chroma and lightness components, respectively. The gray dashed lines indicate the line of mean slope. The dashed dark green and black lines indicate the 95% confidence intervals for the slopes of the chroma and lightness simulations, respectively. The solid black lines (coincident with gray dashed line for the lightness component) indicate the estimated slopes from the original 12 adult data sets.

these as reduced data sets. Then, we sampled 1,000 sets of 21 data sets with replacement from the 12 reduced data sets. In each of these sets, any of the 12 adult reduced data sets could appear multiple times or not at all. To generate different observers, we generated new random responses for each trial using the estimated choice probabilities from the MLCM fit to the data and Equation 6.

To simulate the random number of trials completed for each infant, we first examined the distribution of number of trials from the infants in our data set (Figure 7a). The distribution looks fairly uniform except at the largest number trials completed. These are the infants that completed all 72 trials or about a quarter of the sample. To approximate this distribution, we created a discrete censored uniform distribution (DCU) on the interval (13, 92) with all data above 72 censored to this value. The minimum and censored values correspond to minimum and maximum from our data set, and the maximum of 92 is the value that we calculated that would lead to about one quarter of the samples being censored. For comparison, we drew 1,000 samples of size 21 from the DCU, binned them according to the breaks in the histogram of Figure 7a and then averaged the values in each bin. The results are shown in Figure 7b with the standard deviations indicated as error bars. Given the variability in each bin, the distribution is not very different from that of Figure 7a.

'We then used the 1,000 sets of 21 trials completed to subsample trials randomly from the reduced data sets to generate sets of data of 21 observers with the number of trials following a similar distribution to that from the infant data sets. Each of these sets of 21 subsampled data sets was fit with the GLMM model that we fit to the infant data. The slopes of the chroma and lightness components were extracted from each fit and averaged. Figure 7c shows the comparison of average slopes for both components (gray dashed lines) and the slopes fit to the 12 reduced data sets (black solid lines). To visualize the variability across the 1,000 repeats, we plotted the predicted lines from each simulation as thin lines of transparent color. The averages are indicated as dashed gray lines with 95% confidence intervals as dark green (chroma) or black (lightness) dashed lines. The similarity of the average simulated results with those obtained from the reduced data sets shows that neither the low responses obtained from the infant nor the reversed contributions of the salience components can be explained by the smaller and more variable sampling imposed by testing the infants.

DISCUSSION

This study used the psychophysical technique of MLCM to investigate how the dimensions of lightness and chroma contribute to perceptual saliency in infants and adults. Previous studies have demonstrated the effectiveness of this method when adults are instructed to judge pairs on specific stimulus dimensions (Gerardin et al., 2014; Ho et al., 2008; Rogers et al., 2016), but it was still an open question as to whether the methods could be extended to infants. Here, we used eye tracking with infant and adult observers, while they viewed pairs of green stimuli that varied in lightness and chroma to obtain first fixations as a choice response measure. We successfully used MLCM to model the first-look data with a GLMM.

The analyses reveal that chroma and lightness both contribute to looking preference in infants and adults and that an additive model best describes the data in both groups. This indicates that the looking response depends on an additive combination of the underlying response components to lightness and chroma and that there is no interaction between the two components. Infants' looking behavior was primarily predicted by lightness, but there was a small positive contribution of chroma. Infants prefer stimuli that are darker and more chromatic. However, for adults, chroma primarily determined the first-look response, and there was a small positive contribution of lightness. Adults prefer stimuli that are more chromatic and lighter.

Interestingly, in our former study, when adults were instructed to judge which stimulus is greener, higher chroma led to a more positive chroma contribution but higher lightness to a more negative lightness contribution (Rogers et al., 2016), qualitatively similar to the infants' salience responses in the current study. This raises the possibility that the mechanisms engaged by the salience of the stimulus in the infants are the same as those employed by an adult judging the chromatic difference between a pair of stimuli. The results in both cases, however, show that both the chroma and the lightness contribute to these behaviors. In the Rogers et al. (2016) study, when adults were instructed to judge which stimulus was lighter, the lightness component dominated and the response was independent of the chroma response. This is unlike either of the adult or infant response patterns in the current study.

Influence of sampling constraints in infant data

Our analysis of the adult data when sampled such that the number of participants and trials is equivalent to the infant data set illustrates that the differences in infant and adult response cannot be explained by the smaller and more variable sampling that occurs when testing infants. The variability in the number of trials sampled makes it difficult to analyze the individual data sets of the infants. Grouping the observers together, as in Figures 3 and 5, is useful for identifying the trends in the data but of less value for statistical tests because the observer-specific sources of variance are ignored thereby reducing the power of the test less. The mixed-effects model framework presented here provides a good solution in that observer specific sources of variability can be included in the model yielding the expected gains in sensitivity. Even the small data sets can be included in such an analysis as they contribute information about the population estimates (Bates et al., 2014).

Consistency of response

The average contribution of chroma to the choices of infants is about six times lower than that of the adults. This could arise if the infants discriminate the chroma differences less well, but also if they were more variable in their preferences. All such sources of variability will be confounded in the decision noise, ε, of the model. This term is included to take into account the fact that observers do not give the same response to the same stimulus when the differences are small. Thus, inconsistencies in the response patterns are tolerated by the model and expressed as smaller differences between estimated scale values. This cannot explain, however, the reversal in salience preferences of the infants that would appear to be a real difference in the patterns of infant and adult preferences.

Development of dimensional interaction

Previous researchers have theorized that perceptual dimensions become more separated over development and that infants process stimuli more holistically than adults do (Kemler, 1983; Kemler & Smith, 1978). However, our findings do not support the idea that there are differences between infants and adults in the extent of dimensional separation for color, as the additive model best fit both infant and adult data. Both data sets are accounted for by a pair of invariant response functions, signaling lightness, and chroma that simply add to generate a salience estimate. The presence of a significant interaction would have supported a more holistic processing in that the component responses would not be able to be disentangled. However, there was a difference between adults and infants in the weights applied to lightness and chroma contributions to the response. There are a number of possible explanation for these differences.

Visual strategy

First, the visual strategy may be different between adults and infants. We theorize that adults are focusing more on the abstract stimulus property of color, and ignoring the background, whereas infants are focusing on the whole screen in a trial, including the background. This may lead to luminance contrast driving infant-looking behavior, which would bias the infants toward lightness-based responses. Previous work has shown that for achromatic stimuli, 4-month-old infants looking preferences are governed by luminance contrast (stimulus to background luminance ratio; Chien, Palmer, & Teller, 2005). Here, lightness and chroma levels were perceptually equated using difference scaling data from adults (Rogers et al., 2016), and the fact that the average responses vary approximately linearly as a function of the stimulus index suggests that this scaling was valid. However, all stimuli were darker than the background, and the lightness levels had greater variation from the background than the chroma levels in CIE delta E. Therefore, the pattern of infants' responses may be driven by a preference to stimuli that were most different to the background. If the above "contrast theory" is correct then infants may fixate more on the edges of the stimuli, whereas adults may focus on the center. It would be informative to test this theory by displaying the stimulus pairs for a set amount of time and measuring looking time to each area of the screen.

Cognitive strategy

A second possible explanation for the difference in dimensional contribution between infants and adults is a difference in the cognitive strategy between the two groups. Adult participants were given instructions for the task ("look at the patch that stands out more, or that most grabs your attention"). With these instructions, we aimed to access the same outcome we measured in infant participants: salience. However, the instruction is vague. By making the salience outcome measure in adults explicit, we may have inadvertently introduced unwanted cognitive strategies. We may be accessing our target measurement of "salience" more successfully in infants than in adults. This theory is supported by the large variation in adult responses (see error bars of Figure 6), which may indicate different interpretations of the task among adult observers. In comparison, the variation among infants is much smaller. This is despite the fact that adults performed many more trials than infants (240 for adults compared to a maximum of 72 for infants). Adult eye-movement strategies are highly dependent on the task or instruction given (Buswell, 1935). It would be informative to examine how the dimensions of color interact to inform looking behavior in adults with a range of different instructions.

Dimensional differences

A final explanation for the difference in results is that there are differences in the perceptual weighting of the dimensions between infants and adults. In 6-month infants, hue preference curves are highly similar at different levels of lightness (Brown & Lindsey, 2013). Furthermore, when color and luminance were equated for visual salience, 6.5-month infants noticed a color change, but not a luminance change when tracking an occluded object (Kaldy et al., 2006). This may lead to the inference that lightness does not have an effect on infants' looking preference. However, in this study, we find that in fact lightness makes a stronger contribution than chroma to looking preference in infants.

The observation in this study that lightness influences infant preference more than chroma may be due to the relative faster maturation of the magnocellular pathway, compared to the parvocellular pathway, in the first year of life (Hammarrenger et al., 2003). Neurons in the magnocellular pathway are more sensitive to luminance contrast, whereas neurons in the parvocellular pathway are more sensitive to red/green chromatic contrast (Lee, Pokorny, Martin, Valbergt, & Smith, 1990; Smith, Pokorny, Davis, & Yeh, 1995). Three- and 4-month infants are more sensitive to luminance contrast than chromatic contrast (Dobkins, Anderson, & Lia, 1999). There is evidence of adult-like performance of the magnocellular pathway in 4-month infants, whereas the parvocellular pathway had not fully developed by this age (Dobkins et al., 1999). Our results may be explained by greater sensitivity to lightness differences than chroma differences in our 6-month infant sample, due to greater magnocellular maturity than parvocellular maturity. Our task does not aim to measure the low-level detectability of chroma and lightness in infants. We do not measure just-noticeable differences in these dimensions in isolation; we measure the visual salience of suprathreshold stimuli that vary on both dimensions. However, infants' low-level perceptual sensitivity for each dimension could still contribute to the visual salience of each dimension on our task. This relationship between detectability and visual salience in infancy deserves further investigation (see Kibbe et al., 2018 for further discussion of this issue).

Future directions

This study has demonstrated that a simple method in combination with statistical modeling based on a signal detection model allows scaling of the contributions of stimulus dimensions in perceptual salience in human infants. The method, like most in developmental science, is limited by the patience of the infant participants. In this study, we successfully modeled three levels of lightness and three levels of chroma. Additional levels or dimensions would likely have resulted in too many trials for infants to complete in one session. Importantly, the use of many infant observers in the framework of mixed models allowed us to overcome the low number of trials recoverable from individual infants. An advantage of the method is that it requires a binary measurement; therefore, we were able to exploit first-look responses (left or right stimulus). This means that the trials moved along rapidly. Other measures of looking time, for example, stimulus fixated for the longest duration, or any other measure that can be used to assign a choice between stimuli (e.g., which stimulus does the infant grab first), could equally be exploited. Once the choices are assigned, the analysis proceeds identically to that presented in the current study. Further studies which compare different measures for the same stimulus set would be of interest.

Maximum likelihood conjoint measurement modeling assumes that the observer is making a choice, and this is the basis for estimation of the decision variable. For example, in Rogers et al. (2016), adult observers viewed pairs of stimuli that varied in lightness and chroma, and made a judgment about whether the left or right stimulus was lighter or more chromatic in different sessions. In this study, we use eye-movement data as a proxy measurement for stimulus choice. It is always challenging to determine exactly what is being measured in infant eye-tracking studies (Aslin, 2007). However, eye movements are widely interpreted as choices in infants across developmental psychology (Civan, Teller, & Palmer, 2005; Teller, 1979).

The results reveal interesting differences in responses to color between adults and infants, and there is now potential for this method to be applied in a wider range of contexts. This method could be used to study dissociations between visual domains over the course of development (Dobkins, 2009) by pairing dimensions from different modalities such as color, motion, form, and depth. The development of early cross-modal correspondences could also be studied in this way. For example, auditory-visual correspondences such as associations between high frequency, small size, and bright color (as in Haryu & Kajikawa, 2012).

CONCLUSIONS

This study demonstrates how a signal detection method can be used to investigate salience and interaction of dimensions in infancy. The contributions of luminance and chroma to salience judgments in infants resembled the contributions of these dimensions to adults judging stimuli on the basis of chroma differences. This study paves the way for future work that aims to understand the contribution of various perceptual dimensions to perceptual and cognitive development to benefit from this method.

ACKNOWLEDGMENTS

We thank the adult participants, and the infant participants and their parents. This research was funded by a European Research Council Grant project, "CATEGORIES," 283605 to AF. KK was supported by the following grants: LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR), ANR-11-BSV4-501, CORE-NETS, ANR-14-CE13-0033, ARCHI-CORE, ANR-15-CE32-0016, CORNET, ANR-17-NEUC-0004, A2P2MC, ANR-17-HBPR-0003, CORTICITY.

AUTHOR CONTRIBUTIONS

AF, MR, and KK conceived the project; MR collected the data; MR and KK analyzed the data; AF, MR, and KK wrote the article; and AF obtained funding.

CONFLICT OF INTERESTS

There are no conflict of interests.

REFERENCES

Abbatecola, C., Gerardin, P., Knoblauch, K., & Kennedy, H. (2016). Face and voice contributions to gender discrimination. *Perception ECVP*, *45*, 331.

Albert, A., & Anderson, J. A. (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika*, *71*(1), 1.

Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*(1), 48–53.

Banks, M. S., & Bennett, P. J. (1988). Optical and photoreceptor immaturities limit the spatial and chromatic vision of human neonates. *Journal of the Optical Society of America A*, *5*(12), 2059.

Banks, M. S., & Salapatek, P. (1981). Infant pattern vision: A new approach based on the contrast sensitivity function. *Journal of Experimental Child Psychology*, *31*(1), 1–45.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7.

Bornstein, M. H. (1985). Colour-name versus shape-name learning in young children. *Journal of Child Language*, *12*(2), 387–393.

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.

Burns, B., & Shepp, B. E. (1988). Dimensional interactions and the structure of psychological space: The representation of hue, saturation, and brightness. *Perception & Psychophysics*, *43*(5), 494–507.

Buswell, G. T. (1935). *How people look at pictures: A study of the psychology and perception in art*. Chicago, IL: University of Chicago Press.

Chien, S. H.-L., Palmer, J., & Teller, D. Y. (2005). Achromatic contrast effects in infants: Adults and 4-month-old infants show similar deviations from Wallach's ratio rule. *Vision Research*, *45*(22), 2854–2861.

Civan, A., Teller, D. Y., & Palmer, J. (2005). Relations among spontaneous preferences, familiarized preferences, and novelty effects: Measurements with forced-choice techniques. *Infancy*, *7*(2), 111–142.

Diesendruck, G., & Bloom, P. (2003). How specific is the shape bias? *Child Development*, *74*(1), 168–178.

Diesendruck, G., Markson, L., & Bloom, P. (2003). Children's reliance on creator's intent in extending names for artifacts. *Psychological Science*, *14*(2), 164–169.

Dobkins, K. R. (2009). Does visual modularity increase over the course of development? *Optometry and Vision Science*, *86*(6), E583–E588.

Dobkins, K. R., Anderson, C. M., & Lia, B. (1999). Infant temporal contrast sensitivity functions (tCSFs) mature earlier for luminance than for chromatic stimuli: Evidence for precocious magnocellular development? *Vision Research*, *39*(19), 3223–3239.

Franklin, A., Bevis, L., Ling, Y., & Hurlbert, A. (2010). Biological components of colour preference in infancy. *Developmental Science*, *13*(2), 346–354.

Franklin, A., Pitchford, N., Hart, L., Davies, I. R. L., Clausse, S., & Jennings, S. (2008). Salience of primary and secondary colours in infancy. *British Journal of Developmental Psychology*, *26*(4), 471–483.

Gerardin, P., Devinck, F., Dojat, M., & Knoblauch, K. (2014). Contributions of contour frequency, amplitude, and luminance to the watercolor effect estimated by conjoint measurement. *Journal of Vision*, *14*(4), 9.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.

Hammarrenger, B., Leporé, F., Lippé, S., Labrosse, M., Guillemot, J. P., & Roy, M. S. (2003). Magnocellular and parvocellular developmental course in infants during the first year of life. *Documenta Ophthalmologica*, *107*(3), 225–233.

Hansmann-Roth, S., & Mamassian, P. (2017). A glossy simultaneous contrast: Conjoint measurements of gloss and lightness. *Iperception*, *8*(1), 204166951668777.

Haryu, E., & Kajikawa, S. (2012). Are higher-frequency sounds brighter in color and smaller in size? Auditory-visual correspondences in 10-month-old infants. *Infant Behavior and Development*, *35*(4), 727–732.

Ho, Y. X., Landy, M. S., & Maloney, L. T. (2008). Conjoint measurement of gloss and surface texture. *Psychological Science*, *19*(2), 196–204.

Indow, T., & Kanazawa, K. (1960). Multidimensional mapping of Munsell colors varying in hue, chroma, and value. *Journal of Experimental Psychology*, *59*(5), 330–336.

Ishihara, S. (2010). *Ishihara's test chart for colour deficiency*. Tokyo, Japan: Kanehara Trading Inc.

Kaldy, Z., & Blaser, E. (2009). How to compare apples and oranges: Infants' object identification tested with equally salient shape, luminance and color changes. *Infancy*, *14*(2), 222–243.

Kaldy, Z., & Blaser, E. (2013). Red to green or fast to slow? Infants' visual working memory for "just salient differences". *Child Development*, *84*(6), 1855–1862.

Kaldy, Z., Blaser, E. A., & Leslie, A. M. (2006). A new method for calibrating perceptual salience across dimensions in infants: The case of color vs. luminance. *Developmental Science*, *9*(5), 482–489.

Kaldy, Z., & Leslie, A. M. (2003). Identification of objects in 9-month-old infants: Integrating "what" and "where" information. *Developmental Science*, *6*(3), 360–373.

Kemler, D. G. (1983). Exploring and reexploring issues of integrality, perceptual sensitivity, and dimensional salience. *Journal of Experimental Child Psychology*, *36*(3), 365–379.

Kemler, D. G., & Smith, L. B. (1978). Is there a developmental trend from integrality to separability in perception? *Journal of Experimental Child Psychology*, *26*(3), 498–507.

Kemler, D. G., & Smith, L. B. (1979). Accessing similarity and dimensional relations: Effects of integrality and separability on the discovery of complex concepts. *Journal of Experimental Psychology: General*, *108*(2), 133–150.

Kibbe, M. M., Kàldy, Z., & Blaser, E. (2018). Rules infants look by: Testing the assumption of transitivity in visual salience. *Infancy*, *23*(2), 156–172.

Knoblauch, K., Bieber, M. L., & Werner, J. S. (1998). M- and L-cones in early infancy: I. VEP responses to receptor-isolating stimuli at 4- and 8-weeks of age. *Vision Research*, *38*(12), 1753–1764.

Knoblauch, K., & Maloney, L. T. (2008). MLDS: Maximum likelihood difference scaling in R. *Journal of Statistical Software*, *25*(2), 1–26.

Knoblauch, K., & Maloney, L. T. (2012). *Modeling psychophysical data in R*. New York, NY: Springer.

Knoblauch, K., & Maloney, L. T. (2014). MLCM: Maximum likelihood conjoint measurement. R package version 0.4.1.

Knoblauch, K., Vital-Durand, F., & Barbur, J. L. (2001). Variation of chromatic sensitivity across the life span. *Vision Research*, *41*(1), 23–36.

Komarova, N. L., & Jameson, K. A. (2013). A quantitative theory of human color choices. *PLoS ONE*, *8*(2), e55986.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*(3), 299–321.

Lanthony, P. (1998). *Album Tritan* (2nd ed). Paris, France: Laboratoire de la Vision des Couleurs.

Lee, B. B., Pokorny, J., Martin, P. R., Valbergt, A., & Smith, V. C. (1990). Luminance and chromatic modulation sensitivity of macaque ganglion cells and human observers. *Journal of the Optical Society of America A*, *7*(12), 2223.

Maloney, L. T., & Yang, J. N. (2003). Maximum likelihood difference scaling. *Journal of Vision*, *3*(8), 5.

McCullagh, P., & Nelder, J. (1989). *Generalized linear models*. London, UK: Chapman & Hall.

Morrone, M., Burr, D., & Fiorentini, A. (1990). Development of contrast sensitivity and acuity of the infant colour system. *Proceedings of the Royal Society B: Biological Sciences*, *242*(1304), 134–139.

Morrone, M., Burr, D., & Fiorentini, A. (1993). Development of infant contrast sensitivity to chromatic stimuli. *Vision Research*, *33*(17), 2535–2552.

Moscatelli, A., Mezzetti, M., & Lacquaniti, F. (2012). Modeling psychophysical data at the population-level: The generalized linear mixed model. *Journal of Vision*, *12*(11), 26.

Nichiporuk, N., Knoblauch, K., Abbatecola, C., & Shevell, S. (2017). The lightness distortion effect: Additive conjoint measurement shows race has a larger influence on perceived lightness of upright than inverted faces. *Journal of Vision*, *17*(10), 245.

Poynton, C. (2012). *Digital video and HD: Algorithms and interfaces*. New York, NY: Elsevier.

Qi, L., Chantler, M. J., Siebert, J. P., & Dong, J. (2015). The joint effect of mesoscale and microscale roughness on perceived gloss. *Vision Research*, *115*, 209–217.

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rogers, M., Knoblauch, K., & Franklin, A. (2016). Maximum likelihood conjoint measurement of lightness and chroma. *Journal of the Optical Society of America A*, *33*(3), A184.

Skelton, A., Catchpole, G., Abbott, J., & Franklin, A. (2017). Biological origins of color categorization. *Proceedings of the National Academy of Sciences*, *114*, 5545–5550.

Smith, L. B. (1979). Perceptual development and category generalization. *Child Development*, *50*(3), 705.

Smith, L. B. (1983). Development of classification: The use of similarity and dimensional relations. *Journal of Experimental Child Psychology*, *36*(1), 150–178.

Smith, L. B., & Kemler, D. G. (1978). Levels of experienced dimensionality in children and adults. *Cognitive Psychology*, *10*(4), 502–532.

Smith, V. C., Pokorny, J., Davis, M., & Yeh, T. (1995). Mechanisms subserving temporal modulation sensitivity in silent-cone substitution. *Journal of the Optical Society of America A*, *12*(2), 241.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149.

Teller, D. Y. (1979). The forced-choice preferential looking procedure: A psychophysical technique for use with human infants. *Infant Behavior and Development*, *2*, 135–153.

Volbrecht, V. J., & Werner, J. S. (1987). Isolation of short-wavelength-sensitive cone photoreceptors in 4-6-week-old human infants. *Vision Research*, *27*(3), 469–478.

Wass, S. V., Smith, T. J., & Johnson, M. H. (2013). Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults. *Behavior Research Methods*, *45*(1), 229–250.

Wilcox, T. (1999). Object individuation: Infants' use of shape, size, pattern, and color. *Cognition*, *72*(2), 125–166.

Wood, S. N. (2015). *Core statistics*. Cambridge, UK: Cambridge University Press.

Woods, R. J., & Wilcox, T. (2010). Covariation of color and luminance facilitate object individuation in infancy. *Developmental Psychology*, *46*(3), 681–690.

Wyszecki, G., & Stiles, W. S. (2000). *Color science: Concepts and methods, quantitative data and formulae* (2nd ed.). New York, NY: Wiley.

Zemach, I., Chang, S., & Teller, D. Y. (2007). Infant color vision: Prediction of infants' spontaneous color preferences. *Vision Research*, *47*(10), 1368–1381.