

# Rules Infants Look By: Testing the Assumption of Transitivity in Visual Salience

Melissa M. Kibbe

*Department of Psychological & Brain Sciences  
Boston University*

Zsuzsa Kàldy, and Erik Blaser

*Department of Psychology  
University of Massachusetts Boston*

What drives infants' attention in complex visual scenes? Early models of infant attention suggested that the degree to which different visual features were detectable determines their attentional priority. Here, we tested this by asking whether two targets—defined by different features, but each equally salient when evaluated independently—would drive attention equally when pitted head-to-head. In Experiment 1, we presented 6-month-old infants with an array of Gabor patches in which a target region varied either in color or spatial frequency from the background. Using a forced-choice preferential-looking method, we measured how readily infants fixated the target as its featural difference from the background was parametrically increased. Then, in Experiment 2, we used these psychometric preference functions to choose values for color and spatial frequency targets that were equally salient (preferred), and pitted them against each other within the same display. We reasoned that if salience is transitive, then the stimuli should be iso-salient and infants should therefore show no systematic preference for either stimulus. On the contrary, we found that infants consistently preferred the color-defined stimulus. This suggests that computing visual salience in more complex scenes needs to include factors above and beyond local salience values.

How do infants decide what to pay attention to in a cluttered scene? Research over the past 50 years has explored how the adult attentional system assigns priority values to different regions. Targets for eye movements and visual attention are chosen in part by how conspicuous, or *salient*, they are (Itti, 2005; Itti & Koch, 2001; Treisman, 1985). Bottom-up, image-based models of human adult visual salience assign higher priority to regions that are, for instance, higher in contrast, more dynamic, or uniquely colored relative to other features or regions (Borji & Itti, 2013; Egeth & Yantis, 1997;

Nothdurft, 1992). These models have had success at predicting gaze behavior in adults, particularly when top-down considerations are minimized, that is, when there is no task (e.g., “find the red target”), prior history of reward associated with a given feature (Anderson, Laurent, & Yantis, 2011; Awh, Belopolsky, & Theeuwes, 2012), or other prevailing goals (Tatler, Hayhoe, Land, & Ballard, 2011).

While progress has been made in measuring the bottom-up determinants of attentional salience in adults and older children (Kooiker, van der Steen, & Pel, 2016), relatively little is known about the basic heuristics that determine visual salience in infancy. And, infants are arguably an easier population to study, as top-down control mechanisms in infants are relatively immature (see e.g., Amso, Haas, & Markant, 2014; Colombo, 2001). Banks and Salapatek (1981) proposed a model where the priority value assigned to a target region is determined by its discriminability. For example, targets that are more readily discriminable based on, say, the infant’s contrast sensitivity function are more likely to be attended and fixated, even if all targets are above infant’s detection threshold. In previous developmental work, we expanded on this idea and used it as a basis for calibrating the salience of visual stimuli. We measured the likelihood that a target would be fixated by an infant, as a function of its parametrically varied feature contrast against some background context (e.g., a stimulus that captures infants’ gaze 75% of the time is deemed more salient than one that is fixated only 60% of the time; Kaldy, Blaser, & Leslie, 2006; Kaldy & Blaser, 2009). This method allowed us to select stimuli with known salience values (i.e., one *Just Salient Difference* (JSD) from a common background) for use in subsequent experiments, such as measuring which of two iso-salient features is more readily stored in visual working memory (Kaldy & Blaser, 2013). We employ a similar technique in Experiment 1 of this study.

However, it remains unclear to what extent, in infancy, the salience of a stimulus defined by one feature dimension can be compared to the salience of another stimulus, defined by a different feature dimension, when placed in direct competition. In this study, we found a color-defined stimulus and a spatial frequency-defined stimulus that were both equally salient when presented independently against the same background—each known to attract infants gaze ~65% of the time within 2 sec. We then asked whether these stimuli would attract attention equally when pitted head-to-head. If visual salience is *transitive* in infancy, these stimuli should draw infants’ attention equally when presented together. However, if visual salience is *intransitive*, one of the feature dimensions may be systematically preferred over the other. Answering this question can provide critical insights into the rules driving the way attention is deployed in infancy, and can help constrain models of both infant and adult salience determination.

### Is salience transitive?

To say that some relationship is *transitive* means that when it relates a to b, and b to c, it also relates a to c. Addition is transitive, so if  $a > b$  and  $c > b$ , then  $c > a$ , or if  $a = b + 4$  and  $c = b + 4$ , then  $a = c$ . Similarly, length is transitive, so if Alysha is a foot taller than Betty, and Crystal is a foot taller than Betty, then Alysha and Crystal should be of equal height when placed back-to-back. Treating salience as a unidimensional relationship can lead to the assumption that salience too is transitive, that is, that the salience of individual targets against a common background and/or with

respect to some common reference point will determine the salience of these targets relative to one another.

Previous work with adults suggests that a stimulus presented in close spatial proximity to a target stimulus from the same dimension (e.g., orientation) may modulate the salience of that target (e.g., Nothdurft, 2000), but little interaction is predicted between stimuli from different dimensions (Itti & Koch, 2001; Nothdurft, 2000), and transitivity has been at least implicitly assumed in most computational modeling work (Borji & Itti, 2013; Itti & Koch, 2001). The Itti and Koch model underlies Saliency Toolbox, a popular tool used in infant research to estimate visual salience values (e.g., Amso et al., 2014; Gluckman & Johnson, 2013; Kwon et al., 2016). While it does not necessarily detract from the impact of the results, in this work (and in our previous work), there has been at least the implicit assumption that salience is transitive. But this does not have to be the case; after all, not all relationships are transitive. Someone might prefer an ice cream sundae over a lone scoop of vanilla, and may also prefer more scoops of ice cream to fewer, yet may consistently choose a sundae over a heap of unadorned vanilla. Transitivity is often an empirical question, and violations are well documented in the decision-making literature in humans (Tversky, 1969) and animals (McNamara, Trimmer, & Houston, 2014).

### Testing for salience transitivity in infants

In this study, we tested 6-month-old infants, whose attention is primarily driven by exogenous factors (Colombo, 2001). First, in Experiment 1, we measured the visual salience of stimuli that varied systematically in either spatial frequency or color saturation. We chose to measure salience of color and spatial frequency, as these features are readily detected by infants (Banks & Ginsburg, 1985; Teller, 1998) and are components of natural scenes. As well, there is thought to be considerable overlap in the mechanisms that determine salience of these dimensions (Nothdurft, 2000), making this an especially strong test for intransitivity.<sup>1</sup>

In Experiment 1, we showed infants arrays of Gabor patches presented on a computer screen. We varied the features of a small,  $3 \times 4$  rectangular region of Gabors by changing either their spatial frequency or their red saturation relative to the background. To measure the perceptual salience of the color- and spatial frequency-defined target regions, following stimulus onset, we tracked whether infants' first saccade was to the target region (forced-choice preferential looking, see Teller, 1979). Visual salience measures were obtained for both color and spatial frequency across a range of stimulus intensities.

---

<sup>1</sup>Nothdurft (2000) tested the overlap of the mechanisms that link feature contrast along various dimensions to salience by testing their additivity, in adults. Making a stimulus stand out by virtue of both its brightness and motion will make it more salient than either of those dimensions alone, but ultimately, less than would be predicted by an additive combination. Thus, how close one gets to an additive combination for a pair of features is a measure of the overlap of the underlying mechanisms that determine salience; a fully additive relationship implies independent mechanisms, while a subadditive combination indicates overlap. Their results showed the greatest overlap when color was one of the feature dimensions, especially the combination of color with orientation. If the overlap was total, then transitivity would be all but guaranteed. So, our use of color and spatial frequency makes for an especially strong test. If we show evidence of intransitivity with stimuli that tap into overlapping mechanisms, intransitivity should be even more likely between other, more independent, dimensions.

Next, in Experiment 2, we used the results of Experiment 1 to ask whether stimuli that were equally salient when presented alone would also drive attention equally when presented together. We selected pairs of color- and spatial frequency-defined Gabors that drove infants' attention equally in Experiment 1, and displayed them together on the same screen. We again measured infants' first saccades following stimulus onset. We predicted that if visual salience is transitive across stimulus dimensions in infancy, then equally salient stimuli should continue to drive infants' attention equally, and thus infants should show no preference for one stimulus over the other. Alternatively, if salience is intransitive, infants should show a systematic preference for one feature over the other, even though the features drove attention equally when presented independently.

## EXPERIMENT 1: SALIENCE CALIBRATION

### Method

#### *Participants*

Fourteen infants participated (mean age: 5 months, 18 days; range: 5 months 1 day to 6 months 20 days; six girls). Thirteen of these infants were full-term. One infant was 6 weeks premature, but otherwise healthy, and for this participant, we used the age corrected for prematurity (corrected age: 6 months 17 days). Two additional infants were tested but excluded due to fussiness. Families with infants were recruited via phone and mailing lists and all infants received a small gift for their participation. This study was conducted according to guidelines laid down in the Declaration of Helsinki. Written informed consent was obtained from a caregiver for each infant prior to data collection. The study was approved by the University of Massachusetts, Boston Institutional Review Board.

#### *Stimuli*

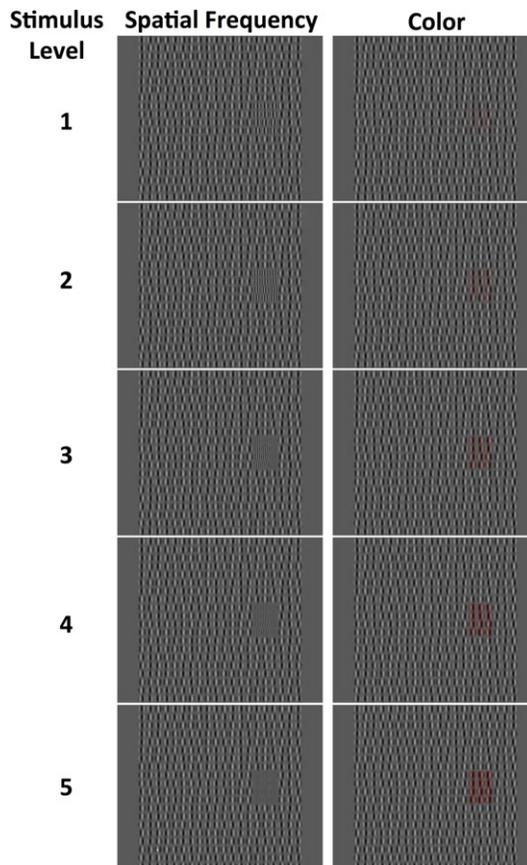
Test stimuli consisted of  $20 \times 20$  arrays of (randomized phase) Gabor<sup>2</sup> patches with spatial frequency of 1 cycle per degree (cpd) and 90% Michelson contrast on a gray ( $125 \text{ cd/m}^2$ ) background. Gabors were chosen because their features can be varied systematically and independently. For example, the spatial frequency of a Gabor can be varied by changing the number of cycles of the sinusoid per visual degree, while the hue can be varied by changing the saturation profile of the luminance modulation. Another advantage of these stimuli is that they roughly reflect the receptive field properties of orientation-selective neurons in the mammalian primary visual cortex (Jones & Palmer, 1987). Gabors have been used extensively as elemental building blocks of stimuli in vision research, including the study of visual salience in adults (e.g., Blaser, Pylyshyn, & Holcombe, 2000; Lee, Itti, Koch, & Braun, 1999), macaques (e.g., McAdams & Maunsell, 2000; Reynolds, Pasternak, & Desimone, 2000), and infants (e.g., Kovács, 2000; Norcia et al., 2005).

---

<sup>2</sup>Gabor functions are sinusoidal gratings of luminance windowed by a 2D Gaussian envelope. Our pseudo-gabor patches were also windowed by a sharp, circular cutoff to enhance visibility and equalize apparent element size.

Within each array, we defined a target region consisting of a  $3 \times 4$  matrix of Gabors, which varied on a single stimulus dimension relative to the background, either in hue saturation or spatial frequency. Hue saturation varied from 6 to 31% red saturation relative to background (0%) in increments of 6.25% (6, 12.25, 18.5, 24.75, 31%). Nominal iso-luminance was maintained by matching lightness values to the background elements ( $\sim 45 L^*$ , in a  $L^*a^*b^*$  space). Spatial frequency varied from 1.5 to 4.5 cpd, in increments of 0.75 cpd (1.5, 2.25, 3, 3.75, 4.5 cpd). Examples of stimuli for each of the five stimulus intensity levels for each target type (spatial frequency or color) are shown in Figure 1. The target region could appear on either the right or left of the array. Between stimulus presentations, a black fixation cross appeared on a gray background to draw infants' attention to the center of the screen. Stimulus onset was accompanied by a female voice announcing the number of the trial in a soft voice, followed by a click sound. Audio was played from a speaker positioned centrally.

Stimuli were created in Matlab and were combined into a movie using Macromedia Director MX. The movie was displayed on a 21" Samsung SyncMaster 213T LCD



**Figure 1** Test displays showing the five stimulus intensity levels for spatial frequency- and color-defined target regions in Experiment 1. The target region could appear on either the left or the right side of the display.

monitor with a refresh rate of 75 Hz. Stimuli are available on Databrary (<https://doi.org/10.17910/b7.357>).

### *Procedure*

Infants sat on their caregiver's lap in a dimly lit, enclosed testing booth. Infants' viewing distance from the test display monitor was approximately 60 cm. Caregivers were instructed to close their eyes for the duration of experiment in order to reduce their influence on infants' behavior.

At the start of the experiment, infants were shown two Familiarization trials designed to familiarize infants with the stimuli and procedure prior to data collection and to give observers an estimate of infants' eye positions relative to the screen. At the start of each Familiarization trial, a fixation cross appeared for 2 sec, followed immediately by a test display with target region of color intensity level of 5 (the maximum intensity value). The display remained visible for 2 sec. In the first Familiarization trial, the target region was on the right; in the second, the target region was on the left. Trials advanced automatically; infants were not required to fixate the cross before the onset of the trial.

Infants were then presented with a series of Salience Calibration trials. Salience Calibration trials proceeded similarly to Familiarization trials. Infants saw a fixation cross, followed by a test display with target region defined by color or spatial frequency presented at one of the five stimulus levels. Each target type (color or spatial frequency at level 1–5) was presented five times, for a total of 50 possible trials. Intensity level, target type (color/spatial frequency), and side of presentation were counter-balanced so that each potential target appeared equally often and in either one of the two locations. Each time a stimulus was presented, it was accompanied by a centrally presented clicking sound to help infants orient to the screen. Figure 2, left panel, shows a sample sequence of two Salience Calibration trials.

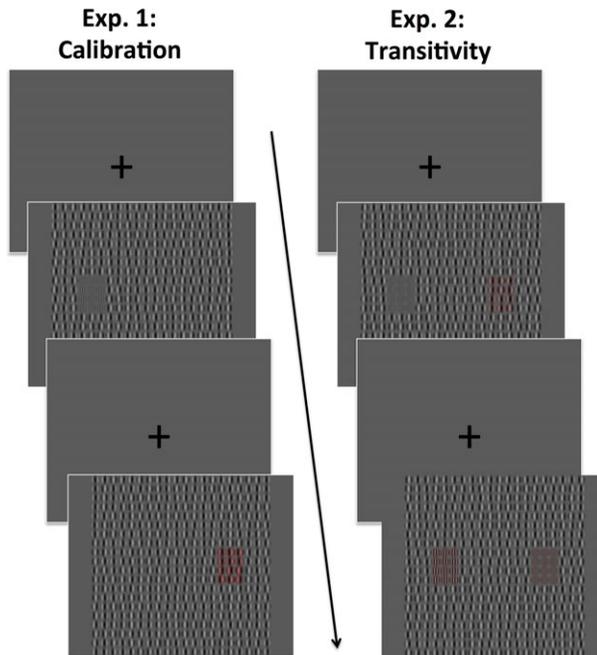
### *Coding*

Infants' gaze was recorded by a camera embedded above the display. Two trained independent observers later viewed the video offline and coded infants' first saccade following stimulus onset for each trial (based on the clicking sound that signaled this event). Observers were naïve to the target type (hue saturation or spatial frequency), level (1–5), and location (left or right) of the target region. Infants' first saccades were coded as "left" or "right." Trials on which infants failed to look at the screen (e.g., because they were distracted) were excluded from analysis. Trials in which observers disagreed about the direction of infants' first saccade were excluded from analysis (mean = 2.7 trials/infant). Infants completed an average of 46.9 trials and gave a valid gaze response (looked to the left or right of the screen) in an average of 32.2 of those trials.

## Results

### *Preference scores*

For each infant, we computed preference scores for each color- and spatial frequency-defined target region at each stimulus intensity level. We calculated the salience

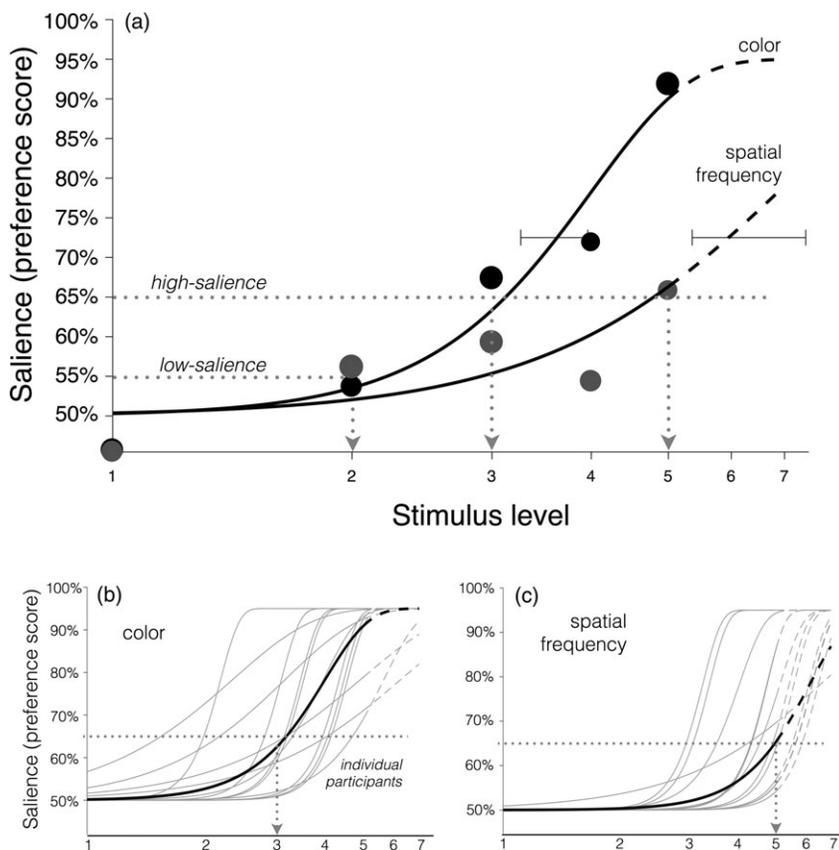


**Figure 2** Left panel shows a sample sequence of two Saliency Calibration trials (Experiment 1). Right panel shows a sample sequence of two Saliency Transitivity trials, high-saliency and hybrid, respectively (Experiment 2).

of each target region by computing the proportion of trials on which the infants' first saccade was to the target region divided by the total number of valid trials. We then averaged infants' preference scores (weighted by the number of valid trials) for each stimulus at each level. Psychometric functions were then fit<sup>3</sup> to these sets of preference scores for each stimulus type (Figure 3a; for reference, curves for each participant for color and spatial frequency are shown in Figure 3b,c, respectively). We performed a goodness-of-fit test using the log-likelihood ratio test ( $G^2$ , based on *deviance*, see Wichmann & Hill, 2001) and found that data corresponded well to the fitted psychometric functions (i.e., we failed to reject the hypothesis that the data were consistent with the specified distributions: color:  $G^2(5) = 1.65$ ,  $p = .89$ ; spatial frequency:  $G^2(5) = 1.81$ ,  $p = .87$ ).

Within the tested range, as stimulus intensity increased, the saliency of the color-defined target regions rose faster than that of the spatial frequency-defined target regions (note, for instance, that infants' preference score for color intensity level 5 was 90%, while their preference for spatial frequency intensity level 5 was 66.3%). This is not unexpected, since, a priori, the scaling between the color and a spatial frequency dimension is undetermined. This difference in infants' preference scores across the two feature dimensions underscores the logic of our experiment: Our method is indifferent to differences in overall saliency across feature values. These functions allowed us to

<sup>3</sup>Weibull fit with guessing rate of 0.50, lapse rate of 0.05, and threshold and slope free to vary [Psignifit 4.0; Schütt, Harmeling, Macke, and Wichmann (2016)].



**Figure 3** Panel a shows infants’ mean preference scores for color- and spatial frequency-defined target regions (black and gray dots, respectively) at each stimulus intensity level (1–5) presented to infants. Psychometric (Weibull) functions for color and spatial frequency are also plotted (dark curve, with dash extrapolation beyond level 5). Error bars show  $\pm 1SE$ . The curves were used to select common low-salience (nominally 55% preference) and high-salience (nominally 65% preference) values for each stimulus type to be used in Experiment 2 (indicated by the dotted gray lines). For reference, panels b and c show curves fitted to individual infants’ preference scores (gray curves) for color (panel b) and spatial frequency (panel c) in addition to curves fitted to the mean of these individual curves (black curves). High-salience values are shown (dotted gray lines).

choose “iso-salient” stimuli for Experiment 2 based on infants’ preferences, not the value of the stimuli themselves.

### Selecting iso-salient stimulus pairs

Using infants’ preference scores, we chose pairs of color- and spatial frequency defined stimuli that were *equally salient* to infants. Using the psychometric functions, we chose two pairs of stimuli: a *low-salience pair* (defined nominally as 55% preference for the target region; actual preference scores were 52.1 and 53.6% for color and spatial frequency, respectively) and a *high-salience pair* (defined nominally as 65%<sup>4</sup>

<sup>4</sup>This is the maximal common saliency level possible in this range, limited by the fact that spatial frequency preference reached a maximum of 66.3% at level 5.

preference for the target region; actual preference scores were 63.4 and 66.3% for color and spatial frequency, respectively). The low-salience pair consisted of color intensity level 2 (12.25% saturation) and spatial frequency level 2 (2.25 cpd). The high-salience pair consisted of color intensity level 3 (18.5% saturation) and spatial frequency level 5 (4.5 cpd). These values were used to define the stimuli used in Experiment 2.

## EXPERIMENT 2: SALIENCE TRANSITIVITY

In Experiment 1, we measured 6-month-old infants' salience functions for color- and spatial frequency-defined stimuli presented separately. The results allowed us to quantify the salience levels of different perceptual dimensions and their ability to independently drive infants' gaze. Next, in Experiment 2, we asked whether stimuli that were equally salient when presented independently would also be equally salient when presented simultaneously. If so, infants' looking preferences should be evenly split between the two stimuli. If, however, salience is not transitive, infants should show a systematic preference for one of the stimulus types.

### Method

#### *Participants*

A separate group of 20 healthy, full-term infants (mean age: 6 months and 2 days; range: 5 months and 2 days–6 months and 27 days; seven girls) participated. Three additional infants were tested but excluded due to fussiness. Families with infants were recruited via phone and mailing lists, and all infants received a small gift for their participation. This study was conducted according to guidelines laid down in the Declaration of Helsinki. Written informed consent was obtained from a caregiver for each infant prior to data collection. The study was approved by the University of Massachusetts, Boston Institutional Review Board.

#### *Stimuli*

As in Experiment 1, stimuli consisted of  $20 \times 20$  arrays of randomized phase Gabor patches (1 cpd) presented on a computer screen. However, unlike in Experiment 1, two target regions were defined, one on the left and one on the right side of the screen. Each target region consisted of a  $3 \times 4$  region of Gabors that varied in either color (hue saturation relative to the background) or spatial frequency (cpd). As described above, in the *low-salience pair*, one target region was presented with color intensity level 2 (12.25% saturation), while the other was presented with spatial frequency intensity level 2 (2.25 cpd). In the *high-salience pair*, one target region presented with color intensity level 3 (18.5% saturation), while the other was presented with spatial frequency level 5 (4.5 cpd). The high-salience pair was our primary test for transitivity. The low-salience pair acted in part as a control, to confirm that stimuli that are insufficient to drive attention in isolation are still disregarded even in multitarget scenes. A third stimulus pair, the *hybrid pair*, was also constructed. The hybrid pair also consisted of two target regions, but each differed along both stimulus dimensions. The high color, low spatial frequency target region was defined by a combination of the

*high-salience* color level and the *low-salience* spatial frequency level. The high spatial frequency, low color target region was defined by the *high-salience* color level and the *low-salience* spatial frequency level. The purpose of including the hybrid pair was to explore whether features with different priority values, combined into a single stimulus, would interact to drive infants' attention differently than when those stimuli were presented alone. Stimuli are available on Databrary (<https://doi.org/10.17910/b7.357>).

The target type (*low-salience*, *high-salience*, or *hybrid*) and the side of placement of each target region in each pair were pseudo-randomized across trials. Each pair was presented 16 times for a total of 48 trials. A black fixation cross on a gray background was shown between trials. Just as in Experiment 1, each time a stimulus was presented, it was accompanied by a centrally presented clicking sound to help infants orient to the screen. Figure 4 shows all three test displays for Experiment 2.

### Procedure

As in Experiment 1, infants were seated on a caregiver's lap in an enclosed testing booth approximately 60 cm from the computer screen. Caregivers were instructed to close their eyes for the duration of experiment.

As in Experiment 1, at the start of the experiment, infants completed two Familiarization trials in order to familiarize them with the stimuli and procedure prior to data collection. Each Familiarization trial consisted of the *high-salience pair*. In one trial, the color target region was presented on the left, and the spatial frequency target region was presented on the right, and in the other trial, the color target was presented on the right and the spatial frequency target on the left. Each 2-sec trial was preceded by a central fixation cross presented for 2 sec.

Immediately following the Familiarization trials, infants observed a series of Salience Transitivity trials. On each trial, infants saw a 2-sec fixation stimulus followed by one of the three stimulus pairs presented for 2 sec. The side of presentation of the color target region (in the case of the low-salience and high-salience pairs) or the higher-value color region (in the case of the hybrid pairs) was determined pseudo-randomly so that targets appeared equally often in either location. Each pair configuration could appear eight times for a potential total of 48 trials. Figure 2, right panel, shows a sample sequence of two Salience Transitivity trials.



**Figure 4** Sample test displays presented in Experiment 2. The side of placement of each target region was varied across trials.

### Coding

All coding was done the same way as in Experiment 1. Infants completed an average of 46.9 trials and had valid gaze responses (looked at the display) on 36.1 of those trials.

### Results

For all stimulus pairs (low-salience, high-salience, hybrid), we calculated each infant's preference for the *color* stimulus by computing the number of trials in which the color stimulus was preferred by the infant divided by the total number of valid trials (trials on which infants looked to either stimulus) for each pair type. For the hybrid pair, we computed infants' preference to look at the side containing the target region with the high-salience color level and the low-salience spatial frequency level. This resulted in a *Color Preference Score* for each infant for each stimulus pair type (low-salience, high-salience, and hybrid). An average Color Preference Score that is significantly different from 50% would suggest that infants preferred to look at one stimulus type over the other. For the low-salience and high-salience pairs, a Color Preference Score significantly >50% would suggest a preference for the color stimulus over the spatial frequency stimulus, while significantly <50% would suggest a preference for spatial frequency over color. For the hybrid pair, a Color Preference Score significantly >50% would suggest a preference for the high-salience color stimulus combined with the low-salience spatial frequency stimulus over the high-salience spatial frequency stimulus combined with the low-salience color stimulus, and a Color Preference Score significantly <50% would suggest the opposite.

We conducted one-sample *t*-tests on Color Preference Scores for each stimulus pair type, comparing infants' average scores to chance level (50%). In addition, we conducted Bayes factor analysis for each stimulus pair. Bayes factor analysis differs from traditional null hypothesis significance testing in that it allows us to either accept or reject the null hypothesis (Gallistel, 2009; Rouder, Speckman, Sun, Morey, & Iverson, 2009). As the null hypothesis—that stimuli that are equally salient independently are also equally salient when presented together—is theoretically interesting, we sought to obtain statistical support for or against this hypothesis. In the case of our results, Bayes factor analysis yields the odds that the samples we observed could come from a distribution with a mean of 50%. A Bayes factor (BF) <3 is considered “weak” evidence for or against the null hypothesis, a BF between 3 and 10 is considered “substantial” evidence, and a BF between 10 and 100 is considered “strong” evidence (Jeffreys, 1961; Gallistel, 2009; Rouder et al., 2009<sup>5</sup>). Finally, we conducted a one-way ANOVA comparing infants' preferences across conditions.

Results are summarized in Table 1. We found that infants' average Color Preference Score was not different from 50% for the low-salience pair (Color Preference Score = 48.1%;  $t_{19} = 0.61$ ,  $p = .55$ , two-tailed; BF = 3.64 in favor of the null), suggesting that infants' first looks were driven equally by both stimuli. However, infants' Color Preference Scores were significantly >50% when infants were presented with the high-salience pair (Color Preference Score = 61.1%;  $t_{19} = 3.37$ ,  $p = .003$ , two-tailed;

---

<sup>5</sup>Bayes Factors were computed using an online Bayes Factor calculator located at <http://pcl.missouri.edu/bayesfactor>.

TABLE 1

Mean Color Preference Scores (Standard Deviation) Averaged Across Infants and Results of One-Sample *T*-Tests for the Three Stimulus Pair Types (Against Chance Level) in Experiment 2

	<i>Mean color preference score (SD)</i>	<i>p (two-tailed)</i>	<i>Cohen's d</i>	<i>Bayes factor</i>
Low-salience pair	48.1% (13.7%)	.55	−0.136	3.64 for null
High-salience pair	61.1% (14.7%)	.003	0.753	13.35 against null
Hybrid pair	60.1% (13.4%)	.003	0.760	14.15 against null

BF = 13.35 against the null), suggesting that infants preferred the color stimulus to the equally salient spatial frequency stimulus. Similarly, infants also preferred the high-salience color stimulus when it was presented in the hybrid pair (Color Preference Score = 60.1%;  $t_{19} = 3.40$ ,  $p = .003$ , two-tailed; BF = 14.15 against the null), suggesting that color drove infants' preferences in the hybrid pair as well.

A one-way ANOVA conducted on infants' preference scores with condition (low-salience, high-salience, hybrid) as a within-subjects factor revealed a significant main effect of condition ( $F_{2,57} = 5.38$ ,  $p = .007$ ). Post hoc tests using Bonferroni correction for multiple comparisons found that infants' preferences in the low-salience condition differed significantly from infants' preferences in both the high-salience ( $p = .014$ ) and hybrid ( $p = .025$ ) conditions, suggesting that infants preferred the color stimulus, as expected, only when the stimulus was above their threshold for detection. However, infants' preferences in the high-salience and hybrid conditions did not differ ( $p = 1.0$ ), suggesting that the color stimulus drove infants' attention similarly in both cases.

Finally, we sought to evaluate the strength of this result, to see how robust it was to (potential) misestimates of the high-salience levels chosen in Experiment 1. If the psychometric functions from Experiment 1, for instance, overestimated color thresholds or underestimated spatial frequency thresholds, this would have led to unbalanced choices (an unfairly high color comparison, or an unfairly low spatial frequency comparison, respectively) heading into Experiment 2. To address this, we first examined the variability associated with the psychometric fit to the color and spatial frequency preference data from Experiment 1 (Figure 3a). For instance, with color, the fit psychometric function had a threshold of color level 3.62, with standard error boundaries of (3.25, 3.97), and 95% CI boundaries of (2.88, 4.31). One way to "simulate" a misestimate, then, is to assume the true color threshold was, for instance,  $-1$  SE from our initial estimate (i.e., 3.25). This would mean that the level 3 color stimulus we used as the high-salience value actually had a salience preference score of 69%<sup>6</sup> (not 63.4% as shown with the original fit at level 3). We then evaluated whether pitting this 69% salience color stimulus against the 66.3% preference score SF stimulus could account for the Color Preference Score bias we observed toward color in Experiment 2. The probability of a bias toward color in this head-to-head comparison can be estimated from these respective preference scores (PS) for color (CL) and spatial frequency (SF):<sup>7</sup>

<sup>6</sup>This preference estimate was derived from a refitting the psychometric function, but with the threshold fixed at the  $-1$  SE threshold value.

<sup>7</sup>In this form, the function is identical to James' "Log 5" (James, 1981), but is comparable to the Bradley-Terry-Luce model (Bradley & Terry, 1952; Luce, 1959).

$$P(\text{CL}, \text{SF}) = \frac{\text{PS}_{\text{CL}}(1 - \text{PS}_{\text{SF}})}{\text{PS}_{\text{CL}}(1 - \text{PS}_{\text{SF}}) + \text{PS}_{\text{SF}}(1 - \text{PS}_{\text{CL}})}$$

Applying this formula predicts that color should “win” the matchup 53% of the time. The actual Color Preference Score we observed of 61.1% (Table 1) exceeds this prediction by a wide margin. Even taking the more extreme scenario of a  $-1$  CI misestimate, implying that a level 3 color stimulus actually had a salience value of 74%, only produces a preference toward color of 59%, again lower than the observed preference score. We repeated this exercise for spatial frequency, using the SE (5.33, 7.50) and CI values (4.74, 8.99) from the preference score fits shown in Figure 3. If spatial frequency thresholds had actually been  $+1$  SE higher, the salience value for spatial frequency used in our high-salience comparison was actually 60%, not 66.3% as initially measured. Pitting this 60% salience value spatial frequency stimulus against the color stimulus with a salience value of 63.4% would produce a preference toward color of 55%, again smaller than what was actually observed. A  $+1$  CI misestimate would yield a preference of 57%, again smaller than the Color Preference Score observed in this experiment. (In fact, even if we simultaneously had *both* overestimated infants’ color thresholds by one SE *and* underestimated infants’ spatial frequency thresholds by one SE, this would not account for the bias observed in Experiment 2, yielding a predicted 59% preference.) This analysis, coupled with the converging evidence from the hybrid pair comparison as well as Bayes factor analysis, indicates that the observed preference for color over spatial frequency was unlikely to have resulted from misestimated salience values, and instead reflects an inherent (and, intransitive) preference for color.

## GENERAL DISCUSSION

In the current study, we asked whether visual salience is *transitive*. Specifically, we asked whether two stimuli, each found independently to have the same salience against a particular background, would be equally salient when pitted head-to-head. Models of infant visual attention (e.g., Banks & Salapatek, 1981) suggest that the salience of a stimulus is determined by the degree to which it is detectable, leading to the prediction that two stimuli that are determined to have the same salience in isolation will drive attention equally when presented together. We tested this prediction in 6-month-old infants, both because we are interested in the rules that guide overt attentional selection during infancy, and because at this age, endogenous attentional systems are relatively immature, meaning attention is largely driven by exogenous factors (Colombo, 2001), thereby reducing the influence of strategies and biases that adults may bring to free-viewing experiments (e.g., Tatler et al., 2011).

In Experiment 1, we first calibrated the perceptual salience of two stimulus dimensions independently: color (red saturation) and spatial frequency (cpd). Then, in Experiment 2, we chose values for these dimensions that were equally salient to infants when presented alone, and pitted them against each other by presenting them simultaneously on the same background. Infants were not given a “task,” and no reward was associated with either feature dimension. We constructed three pairs of stimuli: a *low-salience*

pair consisting of minimally salient targets, a *high-salience* pair consisting of targets that drove infants' looking on a majority of trials, and a *hybrid* pair in which each target consisted of the low-salience value of one stimulus dimension and the high-salience value of the other. We reasoned that if the stimuli were detectable to infants (that is, high-salience) and visual salience is transitive, then infants should look equally at stimuli that were equally salient when presented independently. In fact, we found that salience was *intransitive*: when presented with the high-salience pair, infants systematically preferred the color-defined stimulus to the spatial frequency-defined stimulus. This pattern held for the hybrid pair as well, providing a confirmatory internal replication. The lack of preference with the low-salience pair, and the lack of impact of the weak feature in the hybrid pair both suggest that, not unsurprisingly, stimulus properties below some critical level do not have the power to influence preferential looking, even in more complex scenes.

The results of these experiments suggest that, in infancy, the salience of visual stimuli is not solely determined by the local priority value of the stimuli relative to some background. This suggests that salience is not completely abstracted from the feature identity, but is instead dependent both on the independent priority value of each object and on the *types* of features that constitute them. Our results contrast with the predictions made by infant (Banks & Salapatek, 1981) and adult (e.g., Itti & Koch, 2001) models of bottom-up attentional selection, and suggest that stimuli from different dimensions, when presented together, may interact to produce salience values that are different from the values that are assigned when stimuli are presented alone.

There is an extensive literature on psychophysical, computational, and neurophysiological approaches to visual saliency and how salience contributes to fixation patterns in human adults. More than 60 computational models have been developed since the classic model of Itti, Koch, and Niebur (1998), and their performance has been compared to large data sets of human gaze fixations in natural scenes (see e.g., Borji & Itti, 2013). Beyond feature-based, bottom-up saliency, several of these models emphasize is the effect of top-down guidance (Hayhoe & Ballard, 2005), scene gist (Torralba, Oliva, Castelhano, & Henderson, 2006), and spatial biases in image viewing (Tatler, 2007), among others. Nothdurft (2000) studied the effects of combinations of features on saliency and found that they were additive, however, not necessarily linear in adults. In their recent analysis of the current state of saliency modeling, Bruce, Wloka, Frosst, Rahman, and Tsotsos (2015) provided a list of model constraints that have been identified in the vision literature, and found that "more specific quantitative analysis might be achieved by considering relative salience of two targets along differing feature dimensions, and their point of equivalence." (Bruce et al., 2015, p. 106). Our findings, albeit limited by low sampling rates in infants, contribute to this new direction in the study of attentional guidance.

Why might salience across multiple dimensions be intransitive? More research is required here, but single cell recording evidence from macaques suggests that the salience of some feature dimensions may be processed more slowly than others (Nothdurft, 2000). Stimuli presented alone may elicit the same looking from infants, but if the attentional priority value of one stimulus is assigned faster than the other, that stimulus may win out. Further, while infants' attention is largely driven by exogenous factors at 6 months, infants themselves may nevertheless have top-down predilections, for instance assigning color a higher priority in a multifeature scene due to endogenous factors that are not currently known.

In addition, our results suggest that there are systematic biases in the way attention is deployed in multifeature scenes. In the present case, we found a prioritization of color-defined targets over texture-defined targets: color is assigned a higher priority value than spatial frequency when these features appear together, even though these features were assigned equivalent priority value when presented separately. Why might color win the competition for attention when pitted against an equally salient spatial frequency-defined stimulus? Here again we can only speculate, but one possibility is that infants could more easily form a category of our stimuli with high-salience color values: “red objects.” Infants can perceive color differences in a categorical way as early as 4 months of age (Franklin, Pilling, & Davies, 2005; Franklin et al., 2008). Meanwhile, the high-salience spatial frequency-defined objects did not lend themselves so straightforwardly to a new category (such as “more-stripy objects”).

Our study has some potential limitations. For example, we tested how a group of 6-month-old infants responded to independent stimuli, and used those responses to test whether salience was transitive using a separate group of 6-month-old infants. This method allowed us to collect a large number of trials for each experiment. However, because we looked at separate groups of infants, we are unable to make inferences about the transitivity of salience across featural dimensions within individual infants. It is possible that salience may be variably transitive across individual infants, yielding individual differences in preference profiles that our method was not sensitive to. Future work could explore this possibility.

In sum, our results yield new insights into the fundamental processes involved in attentional selection in infancy. Models of both infant and adult bottom-up visual attention predict that the attentional priority value assigned to features separately defines how those features will drive attention when present together. We found instead that when iso-salient color and spatial frequency stimuli were presented together, infants preferred color. These results contrast with model predictions and suggest that stimuli that drive attention equally when presented independently may interact when presented together, vying for attention. These results may help refine models of salience computation by showing evidence for *intransitivity*; salience computations over scenes with multiple targets may use not just the featural contrast of targets relative to the background, but also higher-level visual considerations.

## ACKNOWLEDGMENTS

This project was supported by National Institutes of Health’s grant R15HD086658 awarded to Z.K. and E.B. The authors would like to thank the help of the undergraduate Research Assistants of the UMass Boston Baby Lab and the families of infant participants who visited the laboratory. The authors declare no conflict of interests in the conduct of this work.

## REFERENCES

- Amso, D., Haas, S., & Markant, J. (2014). An eye tracking investigation of developmental change in bottom-up attention orienting to faces in cluttered natural scenes. *PLoS One*, *9*(1), e85701.

- Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 10367–10371.
- Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*, 437–443.
- Banks, M. S., & Ginsburg, A. P. (1985). Infant visual preferences: A review and new theoretical treatment. *Advances in Child Development and Behavior*, *19*, 207–246.
- 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.
- Blaser, E., Pylyshyn, Z. W., & Holcombe, A. O. (2000). Tracking an object through feature space. *Nature*, *408*, 196–199.
- Borji, A., & Itti, L. (2013). State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(1), 185–207.
- Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, *39*(3/4), 324–345.
- Bruce, N. D. B., Wloka, C., Frosst, N., Rahman, S., & Tsotsos, J. K. (2015). On computational modeling of visual saliency: Examining what's right, and what's left. *Vision Research*, *116*(Pt B), 95–112.
- Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*(1), 337–367.
- Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology*, *48*(1), 269–297.
- Franklin, A., Drivonikou, G. V., Bevis, L., Davies, I. R., Kay, P., & Regier, T. (2008). Categorical perception of color is lateralized to the right hemisphere in infants, but to the left hemisphere in adults. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(9), 3221–3225.
- Franklin, A., Pilling, M., & Davies, I. (2005). The nature of infant color categorization: Evidence from eye movements on a target detection task. *Journal of Experimental Child Psychology*, *91*, 227–248.
- Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review*, *116*(2), 439–453.
- Gluckman, M., & Johnson, S. P. (2013). Attentional capture by social stimuli in young infants. *Frontiers in Psychology*, *4*, 527.
- Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, *9*, 188–194.
- Itti, L. (2005). Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. *Visual Cognition*, *12*, 1093–1123.
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*, 194–203.
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(11), 1254–1259.
- James, B. (1981). Pythagoras and the logarithms. *1981 baseball abstract* (pp. 104–110). Lawrence, KS: Self-Published.
- Jeffreys, H. (1961). *Theory of probability*, 3rd ed. Oxford, UK: Oxford University Press.
- Jones, J. P., & Palmer, L. A. (1987). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, *58*, 1233–1258.
- 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*, 222–243.
- Kaldy, Z., & Blaser, E. (2013). Red to green or fast to slow?: Visual working memory for 'Just Salient Differences'. *Child Development*, *84*, 1855–1862.
- Kaldy, Z., Blaser, E., & Leslie, A. M. (2006). A new method for calibrating perceptual salience across dimensions in infants: The case of color vs. luminance. *Developmental Science*, *9*, 482–489.
- Kooiker, M. J., van der Steen, J., & Pel, J. J. (2016). Development of salience-driven and visually-guided eye movement responses. *Journal of Vision*, *16*(5), 1–11.
- Kovács, I. (2000). Human development of perceptual organization. *Vision Research*, *40*, 1301–1310.
- Kwon, M. K., Setoodehnia, M., Baek, J., Luck, S. J., & Oakes, L. M. (2016). The development of visual search in infancy: Attention to faces versus salience. *Developmental psychology*, *52*(4), 537.
- Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature Neuroscience*, *2*, 375–381.
- Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York, NY: Wiley.
- McAdams, C. J., & Maunsell, J. H. (2000). Attention to both space and feature modulates neuronal responses in macaque area V4. *Journal of Neurophysiology*, *83*(3), 1751–1755.

- McNamara, J. M., Trimmer, P. C., & Houston, A. I. (2014). Natural selection can favour “irrational” behaviour. *Biology Letters*, *10*(1), 20130935.
- Norcia, A. M., Pei, F., Bonneh, Y., Hou, C., Sampath, V., & Pettet, M. W. (2005). Development of sensitivity to texture and contour information in the human infant. *Journal of Cognitive Neuroscience*, *17*, 569–579.
- Nothdurft, H. C. (1992). Feature analysis and the role of similarity in preattentive vision. *Attention, Perception & Psychophysics*, *52*, 355–375.
- Nothdurft, H. C. (2000). Salience from feature contrast: Additivity across dimensions. *Vision Research*, *40*, 1183–1201.
- Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron*, *26*, 703–714.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*, 225–237.
- Schütt, H., Harmeling, S., Macke, J., & Wichmann, F. (2016). Psignifit 4: Pain-free bayesian inference for psychometric functions. *Journal of Vision*, *122*, 105–123.
- Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, *7*(14), 4, 1–17.
- Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision*, *11*(5), 1–23.
- 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.
- Teller, D. Y. (1998). Spatial and temporal aspects of infant color vision. *Vision Research*, *38*, 3275–3282.
- Torralba, A., Oliva, A., Castelano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*, 766–786.
- Treisman, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing*, *31*, 156–177.
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, *76*(1), 31–48.
- Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Attention, Perception & Psychophysics*, *63*, 1293–1313.