Development of salience-driven and visually-guided eye movement responses

Marlou J. G. Kooiker

Vestibular and Ocular Motor Research Group, Department of Neuroscience, Erasmus University Medical Center, Rotterdam, The Netherlands

✉

Johannes van der Steen

Vestibular and Ocular Motor Research Group, Department of Neuroscience, Erasmus University Medical Center, Rotterdam, The Netherlands
Royal Dutch Visio, Centre of Expertise for Blind and Partially Sighted People, Huizen, the Netherlands

✉

Johan J. M. Pel

Vestibular and Ocular Motor Research Group, Department of Neuroscience, Erasmus University Medical Center, Rotterdam, The Netherlands

✉

Development of visuospatial attention can be quantified from infancy onward using visually-guided eye movement responses. We investigated the interaction between eye movement response times and salience in target areas of visual stimuli over age in a cohort of typically developing children. A preferential looking (PL) paradigm consisting of stimuli with six different visual modalities (cartoons, contrast, form, local motion, color, global motion) was combined with the automated measurement of reflexive eye movements. Effective salience was defined as visual salience of each target area relative to its background. Three classes of PL stimuli were used: with high- (cartoon, contrast), intermediate- (local motion, form), and low-effective salience (global motion, color). Eye movement response times to the target areas of the six PL stimuli were nonverbally assessed in 220 children aged 1–12 years. The development of response times with age was influenced by effective salience: Response times to targets with high salience reached stable values earlier in development (around 4 years of age) than to targets with low salience (around 9 years of age). Intra-individual response time variability was highest for low-salient stimuli, and stabilized later (around 4 years) than for highly salient stimuli (2 years). The improvement of eye movement response times to visual modalities in PL stimuli occurred earlier in development for highly salient than for low-salient targets. The present age-dependent and salience-related results provide a quantitative and theoretical framework to assess the development of visuospatial attention, and of related visual processing capacities, in children from 1 year of age.

Introduction

In humans, viewing behavior requires continuous refixations of the eyes, due to the highly specialized retinal organization with a foveal and a peripheral region. Therefore, the allocation of visuospatial attention depends on complex interactions between the oculomotor and the visual system. Eye movements can be directed toward a new location in a visual scene either voluntarily (goal-driven) or reflexively (stimulus-driven; Corbetta, Patel, & Shulman, 2008). A fundamental problem for a long time has been to quantify visuospatial attention. How does one compare visuo-spatial attention for one visual scene with that for another scene? An elegant solution for this problem comes from computational models that quantify the contribution of different visual attributes in images and scenes in so-called salience maps (Itti & Koch, 2000; Koch & Ullman, 1985). Separate feature maps encode spatial contrast based on, for example, colors, intensity, orientation, or motion. These feature maps are combined in a topographical salience map that contains the most conspicuous areas resulting from parallel processing of the visual features. The development of such algorithms has evolved from computing salience

Age group	1	2	3	4	5	6	7	8	9	10	11
N	35	19	36	11	18	23	20	9	10	15	22

Table 1. Number of included children per age group.

of simple, static images to applicability in complex scenes and video sequences (Schütz, Braun, & Gegenfurtner, 2011). Examples are graph algorithms to achieve natural salience computations (Harel, Koch, & Perona, 2006), or models based on information seeking linked to neural control in visual cortex (Bruce & Tsotsos, 2009). In general, salience maps are used to predict fixation locations and have been proved useful to investigate the underlying principles of selective visuospatial attention in adults (Fecteau & Munoz, 2006; Koch & Ullman, 1985; Treue, 2003).

So far, little is known on how visuospatial attention allocation develops in the first years of life. It has been proposed that during childhood, visual exploration of images is strongly affected by salience information (bottom-up features), whereas inner goals (top-down strategies) play a larger role in the guidance of eye movements later during development (Açik, Sarwary, Schultze-Kraft, Onat, & König, 2010). In addition, aspects of the oculomotor system, e.g., latency and accuracy of eye movements, continue to develop during childhood (Luna, Velanova, & Geier, 2008). The developmental timeline of oculomotor control depends on its specific aspect and the extent of top-down involvement. Reflexive, exogenous driven saccadic responses are known to mature earlier in development and along different trajectories than tasks requiring more cognitive control, such as saccadic inhibition or memory-guided saccades, which develop up to adolescence (Karatekin, 2007; Luna, Garver, Urban, Lazar, & Sweeney, 2004; Luna et al., 2008).

A general accepted method to probe a child's ability to detect features in a visual scene is the preferential looking paradigm (PL; Fantz, 1965). By observing if a child makes eye movements toward a visual stimulus, a child's visuospatial attentional performance can be investigated from infancy onwards. Based on the eye movement responses evoked by visual stimuli, conclusions may be drawn about a child's ability to detect visual features such as global motion and form coherences, or the ability to perceive faces (Kanazawa, Shirai, Ohtsuka, & Yamaguchi, 2006; Wattam-Bell, 1996). By combining PL paradigms with remote eye tracking (Pel, Manders, & van der Steen, 2010), we have established normative data on the development of visually-guided viewing behavior in typically developing children (aged 0–12 years). Within this group, we found a consistent hierarchy in response times to different visual modalities. For example, the execution of reflexive eye movements to coherent form- and -

motion stimuli takes about two to three times longer (300–600 ms) than eye movements to oscillating high-contrast (cartoon) stimuli (~200 ms; Boot, Pel, Evenhuis, & van der Steen, 2012; Kooiker, Pel, & van der Steen, 2014). These differences in response times may not only arise from differential processing of the visual modalities, but also from differences in salience between stimuli that are caused by low-level image features. In addition, the effect of stimulus salience on viewing behavior may interact with age-related development of response times.

The aim of the present study was to investigate the relation between visual salience of the target areas in PL stimuli and the development of reflexive eye movement responses in typically developing children from 1 to 12 years of age. Within a commonly used, standardized set of visual PL stimuli (Boot et al., 2012; Kooiker et al., 2014; Pel et al., 2010), we calculated the salience of the target area in each PL stimulus relative to its background. This effective salience value was compared with the corresponding eye movement response times and their variability, measured in different age groups. We hypothesized that the more salient the target area and the older the children, the faster the response times. In addition, we expected that response times for visual stimuli with low target salience reach stable values at a later age.

Methods

Participants

This study is part of a longitudinal project investigating the development of visual processing capacities with an eye tracking-based test paradigm. Visually-guided eye movements were measured in 220 typically developing children between 1 and 12 years old, with a mean age (SD) of 5.7 (3.3) years. Age groups of one year each were constructed. Table 1 shows the number of children in each age group. Children were recruited at daycare centers and included for participation when they had no history of visual impairments, neurological damage, or premature birth. Informed consent was obtained from parents of all included children. Experimental procedures were approved by the Medical Ethical Committee of the Erasmus University Medical Centre, Rotterdam, The Netherlands (METC-2006-055 and MEC 2012-097) and adhered to the tenets
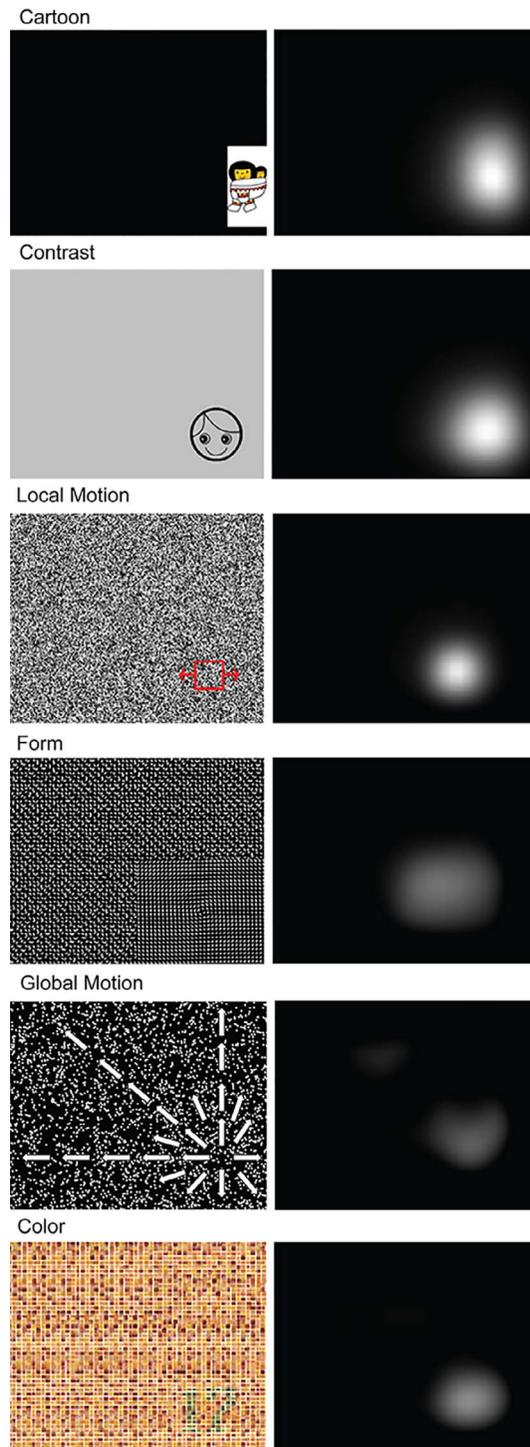
Figure 1. The left column presents an overview of the six visual stimuli that were presented and for which effective salience (ES) was calculated. Each stimulus was a short movie that contained a specific salient area defined as the target area. For illustration purposes, the target area of each stimulus is presented in the lower right corner. The right column presents the corresponding salience map of each stimulus resulting from the presented two-step approach, using the nontarget salience correction (see Figure 2).

of the Declaration of Helsinki (World Medical Association, 2013).

Procedure

Each child sat in a comfortable chair or on the lap of its parent or caregiver at approximately 60 cm distance of an eye tracker monitor (Tobii T60XL, Tobii, Sweden). This system recorded the child's eye movements at a sampling rate of 60 Hz, and compensated free head movements. The visual angle towards the monitor was approximately $30° \times 24°$ (1280 × 1024 pixels), and the system's latency was ± 30 ms. After a standardized five-point calibration procedure, two test sequences containing PL stimuli were shown in random order on the eye tracker monitor. These PL stimuli were designed to test the effectiveness of processing distinct visual modalities and are described in the next section. Each stimulus was presented for 4 s in a target area that was located in one of the four monitor quadrants. All stimuli were presented four times: once in each monitor quadrant. Cartoons were presented 16 times, to retain children's visual attention and enable postcalibration of the data. Note that prior to the presentation of each new stimulus, no fixation point was shown. Stimuli were presented at random, but the location of the target area was always changed from one quadrant to another. This resulted in a minimal visual viewing angle of 18° between the centers of two consecutive target areas. Two test sequences were designed that both contained the same stimuli but in a different randomized order. In total, 192 stimuli were presented (6 PL stimuli × 4 quadrants × 4 times × 2 sequences). In addition, some stimuli were presented to test basic eye movements, such as fixation and pursuit. The total test duration was ~14 minutes.

PL stimuli

To test the effectiveness of processing specific visual modalities in PL stimuli and to compare their salience, three dynamic visual stimuli (oscillating Cartoons, Global coherent motion, and Local motion) and three static visual stimuli (coherent Form, Contrast, and Color) were included (see Figure 1, left column). Their target areas were located in one of four monitor quadrants. These targets were all defined by a distinct visual modality that distinguished them from a background that did not contain that specific modality:

Cartoon: a movie of a colorful, high contrast, slowly oscillating cartoon picture (reproduced with permission from Dick Bruna, Mercis BV, Amsterdam, The Netherlands) with a visual angle of $4.5° \times 9.0°$ (width ×
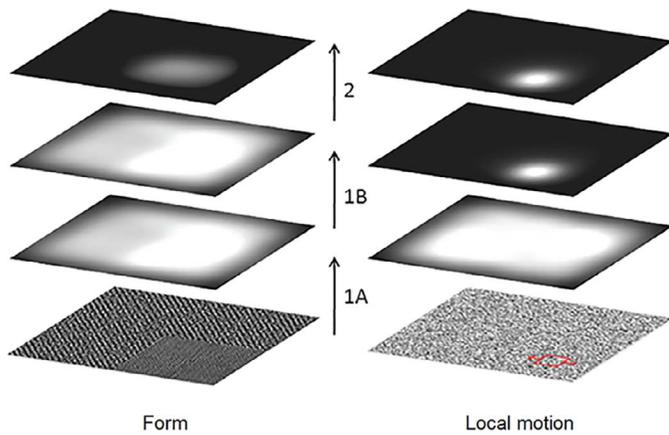
Figure 2. The two-step approach to calculate effective salience, displayed for Form and Local motion. Step 1: calculation of a standard salience map using two consecutive image presentations (step 1A and 1B) and step 2: subtracting the nontarget salience information from the complete salience map. This approach resulted in clear visualization of the effective salience information in the target areas compared to the nontarget areas of stimuli. For Form, the salience of the coherently oriented white lines became visible. For Local motion, salience of the moving target became visible.

height) moving 1.5° up and down at 3°/s, against a black background.

Global coherent motion: a movie containing an array of white dots (diameter 0.25°; density 2.6 dots/degree^2) against a black background. The dots expand from the center of the target area (a quadrant of 8°) to the borders of the monitor at 11.8°/s and with a limited life time of 0.4 s.

Local motion: a movie containing a black-and-white patterned square target, with a visual angle of 2.3°, against an equally patterned background, moving 2.5° to the left and to the right at 2.5 °/s.

Coherent form: an image with an array of randomly oriented short white lines (0.2° × 0.6°; density 4.3 lines/degree^2) against a black background. In the target area (an 8° quadrant), all lines are oriented to form a circle.

Contrast: an image with a 0% brightness (black) picture of a face (adapted from Hiding Heidi, Hyvärinen) was plotted against a 75% brightness (light gray) background. This image is defined by shape and contrast, but is used in clinical practice to assess contrast sensitivity.

Color: an image with a dotted green number 17 in the target area against a red-yellow dotted background.

Effective salience (ES)

The PL stimuli used to test visual information processing had target areas that were either presented against a nonpatterned uniform background (e.g., Cartoon or Contrast stimulus), or against a similarly designed background that lacked only that specific visual modality (e.g., coherent Form-, or Local motion stimulus). To compare the salience of the target areas between stimuli, the salience of their backgrounds was taken into account. We computed the parameter effective salience (ES) to determine the degree of salience within a target area of a stimulus relative to its background. This was done in two steps: by calculating standard salience maps and values for the total stimulus, and then calculating salience of the target area relative to the nontarget background. First, a standard salience map of a PL stimulus was constructed using the graph-based visual salience algorithm (GBVS; Harel, 2011; Harel et al., 2006) for MATLAB (MATLAB, Natick, MA). This algorithm computed salience maps using feature channels like contrast, intensity, and motion, by combining the salient information of preceding maps. In the present study, we used two consecutive images to construct a standard salience map of both static and dynamic stimuli. We activated the channels that corresponded with the visual information content of each visual stimulus' target area. The activated channels represented the visual modalities the target was defined by. The channels orientation, contrast, color, and flicker were activated for Cartoon, whereas only flicker was activated for Local motion and Global motion. The channel orientation was used for Form, color for Color, and contrast and orientation for Contrast. Standard settings for weights of feature channels (i.e., all weights '1') and orientations were applied and the global–mean local maxima scheme was selected as normalization algorithm (Harel, 2011; Harel et al., 2006). Figure 2 shows the two iteration steps to construct the standard salience map for sequences of two image presentations of the stimuli Form and Local motion (1280 × 1024 pixel images), each with the target area in the lower right corner (see step 1A; first image presentation, and 1B; second image presentation, in Figure 2). Simultaneously, the salience value in terms of percentage of salient pixels was calculated. Next, the average salience value in the nontarget area (three quadrants) was subtracted from the salience value in the target area (one quadrant; step 2). This was done by calculating the 85th percentile of the salience value in the nontarget areas and subtracting this value from the corresponding total salience value of the standard salience map. This step was added to calculate the relative contribution of the visual information presented in the target area to overall stimulus salience, thereby assuming a linear relation. In Figure 2 it can be seen that the salience map of Local motion remained equal after step 2, since the 85th percentile salience of the nontarget area was about zero. However, for Form,

Gaze to target (°)

A   Time (s)
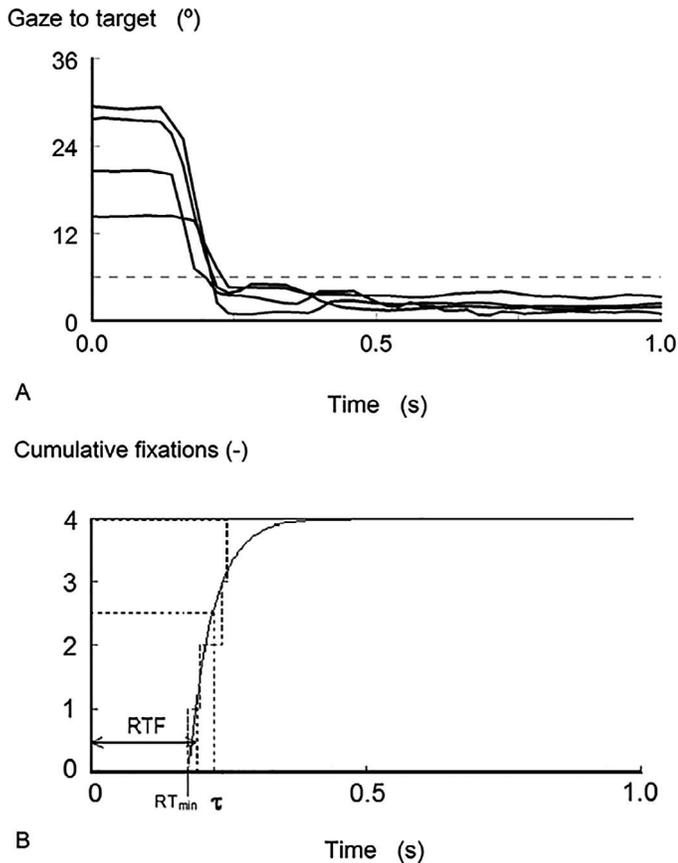
Cumulative fixations (-)

B   Time (s)

Figure 3. Panel A shows four gaze responses of one subject expressed as the distance from the center of the target area (degrees of visual angle, *y*-axis), over time (seconds, *x*-axis). In panel B, the corresponding reaction times of these responses are arranged in a cumulative plot. An exponential curve was fitted to this plot to quantify this subject's reaction time to fixation (RTF): the sum of the minimum reaction time (RT_{min}), and 1/3 of the time constant τ of the exponential curve.

step 2 resulted in a clear contribution of the coherent form to salience in the target area compared to the nontarget areas, i.e., effective salience.

Figure 1 (right column) shows for each PL stimulus the adapted salience map that resulted from the two-step approach. For illustration purposes, the lower right quadrant was depicted as the target area. As a last step, we divided the sum of salience values of all pixels in the target area by the total number of stimulus pixels, multiplied by 100%. This parameter value was a measure for effective salience (ES) in the target area compared to the background, with a maximum value of 25% (given a PL stimulus with four potential target areas). This value was calculated for each of the six PL stimuli used in this study. Highest effective salience values were found for Cartoon (ES = 8.84%), and Contrast (ES = 8.28%). Intermediate effective salience was found for Form (ES = 4.47%), and Local motion

(ES = 4.70%). Lowest effective salience was found for Color (ES = 1.63%), and Global motion (ES = 0.68%).

Reflexive eye movement responses

Reflexive eye movement responses to target areas of the visual stimuli were analyzed by recalculating eye movement data as a visual angle between the gaze location on the monitor and the center of the target area, using the average viewing distance. For each target area, a circular area of interest (AOI) was defined with a radius of 6° for Cartoon, Contrast, Color, and Global motion, and a radius of 8° for Global motion and Form. The AOI for Global motion and Form was chosen slightly larger, because the target areas of both stimuli covered approximately a full quadrant. The center of each AOI corresponded with the center of the target area. The eye movement response time was defined as the time between stimulus presentation and gaze entering the AOI, i.e., Reaction Time (RT; Figure 3). Eye movements to each stimulus presentation were manually analyzed to only include RT values corresponding with reflexive re-sponses to the target area (i.e., within 1 s and not part of a visual search pattern). Since each visual stimulus is shown ≥ four times during a test sequence, a number of RT values per visual stimulus become available at the end of a test (Figure 3A). A cumulative plot was constructed of all available RTs (at least one; Figure 3B). An exponential function was fitted to this plot to quantify the best estimate of the average RT per tested visual modality, represented by the Reaction Time to Fixation (RTF; Pel et al., 2010). RTF is defined as:

$$RTF = RT_{min} + 1/3\tau \qquad (1)$$

In this equation and in Figure 3, RT_{min} reflects the minimum (i.e., fastest) processing time of a specific visual stimulus (including eye movement execution), whereas $1/3 \tau$ reflects 1/3 of the time constant τ of the exponential function. 1/3 of τ is a measure for dispersion in RT values within one subject, represent-ing intra-individual variability in RT. When visual attention or visual information processing varies over time, τ may be prolonged. Thus, RTF is the average time it takes to process the visual information contained by a specific stimulus and to execute an eye movement toward it.

Statistics

Eleven age bins of 1 year each were constructed (i.e., from 1 to 12 years of age; the 11th group contained all 11-, and four 12-year olds). The effects of stimulus type
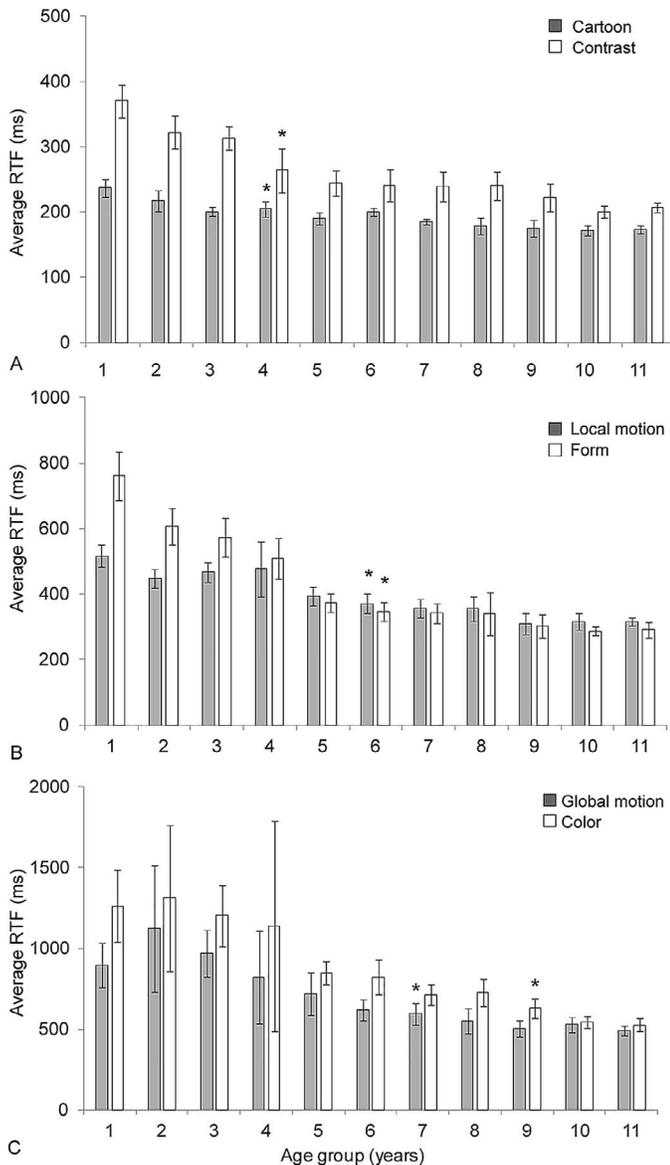
Figure 4. Bar graphs displaying average reaction time to fixation (RTF; ms) as a function of age group, separately for PL stimuli with high ES, i.e., Cartoon and Contrast (panel A), stimuli with intermediate ES, i.e., Local motion and Form (panel B), and stimuli with low ES, i.e., Global motion and Color (panel C). Asterisks represent the age at which RTF reached stable levels. Note the differences in values on the y-axes. Error bars represent standard errors ($\pm$ 2 SE).

and age group on RTF were tested with a mixed analysis of variance (ANOVA). When a main effect of stimulus type was found, differences in RTF between all pairs of stimuli were assessed with paired t tests. The strength of the relation between RTF and ES values of the stimuli was computed with Pearson's correlation coefficients. When a main effect of age group was found, a one-way ANOVA of age group on RTF was done per stimulus. Post-hoc Games–Howell tests were used to determine the age group from which RTF

values did no longer significantly differ from RTF in the oldest age group (i.e., the age at which stable RTF values were reached). One-way ANOVAs per stimulus were also done to determine the effect of age group on RT_{min} and τ. All statistical tests were performed using SPSS version 20.0 (IBM SPSS Statistics, Chicago, IL).

Results

Reflexive eye movement responses (RTF)

The number of children in which stimulus-related eye movement responses were obtained differed between stimuli: Response rates were lower for Color ($N = 161$ of 220) and Global motion ($N = 205$ of 220) than for the other stimuli (all $N \geq 208$ of 220). The number of obtained responses to stimuli correlated significantly with their ES values ($r_p = 0.57$, $p < .05$). Since the assumption of sphericity in the mixed ANOVA was not met [$\chi^2(5) = 899$, $p < 0.01$], the degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity ($\varepsilon = 0.43$). A significant effect of PL stimulus type on RTF was found, $F(2,1) = 205$, $p < 0.001$. Overall, Cartoon triggered the fastest RTF values [mean (SD): 202 (34) ms; and Color the slowest (mean (SD): 908 (454) ms]. In the total group, a strong negative correlation was found between average RTF and ES ($r_p = -0.93$, $p < 0.01$). In addition, age group had a significant effect on RTF values to all stimulus types, $F(10) = 30.8$, $p < 0.001$. Also, a significant interaction between age group and stimulus type on RTF was found, $F(21,4) = 5.41$, $p < 0.001$, indicating that the effect of age differed for the different stimulus types, i.e., ES values. Therefore, development of RTF over age will be reported separately for stimuli with high-, intermediate-, and low ES levels.

Development of RTF for PL stimuli with different salience levels

Figure 4 shows mean RTF values and standard errors (2 SE) over age, separately for stimuli with high ES (Cartoon and Contrast, panel A), intermediate ES (Local motion and Form; panel B), and low ES (Global motion and Color; panel C). For Cartoon and Contrast, stable RTF values were reached at the age of four years. RTF values were significantly lower for the dynamic Cartoon stimulus compared to the static Contrast stimulus (mean difference Cartoon − Contrast = −80 ms; $t(211) = -17.59$, $p < 0.001$). For Local motion and Form, stable values were reached at the age of 6 years. RTF values were significantly lower for the dynamic
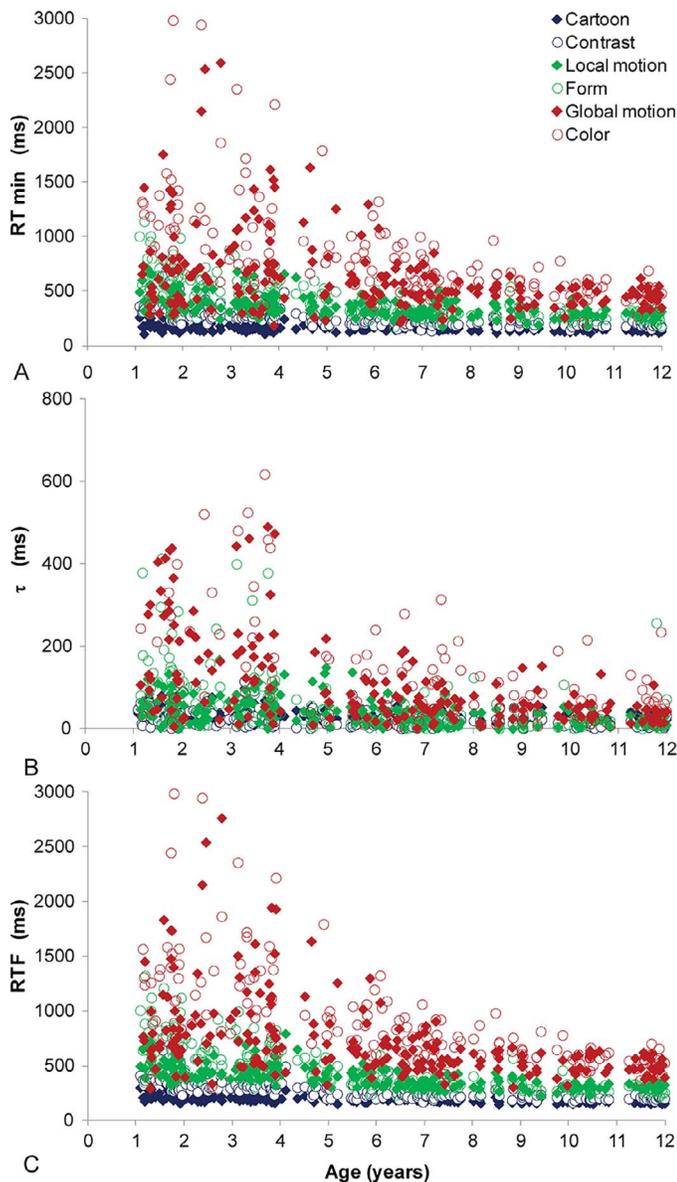
Figure 5. Panel A shows the RT_{min} values (i.e., fastest eye movement reaction time) over age. Panel B shows the τ values (i.e., intra-individual variability in reaction times) over age. Both values are combined per subject and comprise reaction time to fixation (RTF), shown in panel C. Values are shown separately for the six PL stimuli. Note the differences in values on the *y*-axes.

Development of intra-individual RTF variability

Figure 5A shows the RT_{min} values, Figure 5B the 1/3 τ values and Figure 5C the RTF values against age, separately for all tested visual stimuli. The RT_{min} and RTF values strongly correlated ($r_p = 0.90$, $p < 0.001$) and showed similar hierarchies as a function of stimulus and age. From the age of 2 years, the τ values of stimuli with high- and intermediate ES showed constant levels over age. The τ values to stimuli with low ES decreased up to the age of 4 years after which they reached stable levels.

Discussion

In the present study, we showed that the salience level in target areas of visual stimuli (i.e., effective salience) is related to the timing of visually-guided eye movement responses in children. Moreover, effective salience influences the development of these response times and their individual variability over age. Highly salient visual locations triggered faster and less variable responses that improved and reached stable levels earlier in development than low-salient visual locations.

The oculomotor orienting system, i.e., bringing an object of interest onto the fovea for stable viewing, is the first in a series of visual mechanisms that develops in human infancy (Atkinson & Braddick, 2001). At birth, this orienting system is mainly subcortically controlled, and responsible for reflexive eye movements. From around three months of age, control of eye movements is gradually more regulated by the integration of subcortical and cortical (parietal, frontal) circuits. These circuits depend on information from cortical selective modules for different visual modalities, e.g., shape or color, and the binding and segmentation of these modalities (Atkinson & Braddick, 2001). Thus, eye movement responses are to a certain extent dependent on the visual information they are guided to.

In a previous study, the time needed for eye movement execution was on average 60 ms in children aged 1 to 10 years (Pel et al., 2010), whereas in the present study this was on average 48 ms (SD 18 ms). These differences were unrelated to age, indicating that development of the oculomotor system from the age of 1 year did not affect development of reflexive, visually-guided, eye movement responses.

Age-related eye movement responses to the different PL stimuli might reflect different developmental trajectories of brain areas in which salience maps are formed. fMRI evidence has been found for a hierarchy of salience maps in human early visual cortex (V1 to hV4). The maps provided precise topographic infor-

Local motion stimulus than for the static Form stimulus (mean difference Local motion − Form = −69 ms; $t(210) = -5.67$, $p < 0.001$). RTF values of Global motion reached stable levels at the age of 7 years, whereas RTF values of Color reached stable levels at the age of 9 years. RTF values were significantly lower for the dynamic Global motion stimulus than for the static Color stimulus (mean difference Global motion − Color = −151 ms; $t(157) = -5.19$, $p < 0.001$).

mation for successful visual search. It was shown that V1 is mainly responsive to bottom-up signals, V2 mainly to top-down modulations, and that in hV4 bottom-up salience and top-down control converge (Melloni, van Leeuwen, Alink, & Muller, 2012). Alternatively, neural visual system development may have affected responses to different visual modalities. It has been argued that form- and motion-related signals arise from a network of extrastriate areas, but that these processing pathways undergo reorganization from infancy to adulthood (Wattam-Bell et al., 2010). Sensitivity to pattern properties (e.g., coherent form) emerges earlier in cortical development than sensitivity to directional motion, but global integration of these directional signals develops in infancy faster and more robustly than for static patterns (Braddick & Atkinson, 2011). At present, it is not clear to which extent the age-related development of response times is related to salience per se, or to increased sensitivity of the visual system to a particular visual modality. This might be investigated by presenting visual stimuli of a specific visual modality with various salience levels to a cohort of young children. Coupling such a paradigm with neuroimaging techniques may elucidate neural origins of these responses.

The present age-dependent and salience-related development of response times may provide a basis for the assessment of visual processing capacities that result in visuospatial attention allocation, in children at risk of visual abnormalities. The presented quantitative and nonverbal approach is particularly promising for children below the age of 4 years, or for intellectually disabled children.

In accordance with results from a smaller cohort of children (Pel et al., 2010), RTF values to Cartoons reached mature values after the first 3 years of life. For Contrast, a similar developmental trajectory was found. These stimuli both contain face-like features. Attentional orienting to facial patterns is known to be present early in infancy (Glückman & Johnson, 2013). It is possible that the presence of face-like features invokes a form of top-down influence in the form of social preference, which overrules the effect of salience (Glückman & Johnson, 2013). However, responses to Cartoons were faster than those to Contrast, in all age groups. The high ES value for Contrast results from contrast and intensity differences between target and background, whereas the ES value of Cartoon invoked additional color, motion, and flicker differences. Eye movement responses to Cartoon may result from activation in multiple visual pathways and areas at once, increasing the likelihood of quickly detecting at least one visual modality of this stimulus. It is likely that the differential maturational time lines reflect the stimulus' ability to attract visual attention allocation. Besides their ability to predict fixation locations,

salience maps may also be used to predict temporal aspects of viewing behavior (e.g., response times).

The age of 4 years, at which stable RTF values were found for the Cartoon stimulus, was lower than that of previous studies investigating children's reflexive oculomotor responses (see for reviews Karatekin, 2007; Luna et al., 2008). Possible explanations may be the familiarity of stimuli for children and stimulus size. Most of the previous studies used small fixation lights (LEDs) as stimuli, opposed to the more meaningful and larger visual images and movies in our study. However, it is possible that RTF continues to decrease up to 15 years of age, but in a different order of magnitude than reported in previous studies (Karatekin, 2007; Luna et al., 2008).

Motion-based segmentation is of high importance in learning to organize the visual world, since it occurs whenever there is motion of objects or self-motion (Braddick & Atkinson, 2011). It has been suggested that motion-based segmentation of visual scenes is particularly salient in driving an infant's visual attention (Braddick & Atkinson, 2011). Therefore, PL stimuli with motion may have triggered faster responses than stimuli with similar salience but without motion, owing to more natural or functional salience as opposed to visual salience per se. In the guidance of gaze, salience by itself is known to have a decreasing influence when the contribution of high-level visual input, or top-down control, increases (Schütz et al., 2011), for example with extended scene viewing time (Helo, Pannasch, Sirri, & Rämä, 2014). Our findings indicate that the decrease of reaction times with age may be stabilized earlier in development for highly salient- than for low salient simple scenes. These findings relate to those found during complex natural scenes, where the effect of salience on viewing behavior was larger in younger children compared to older ones (Helo et al., 2014).

A large body of literature reported on the development of processing coherent form and coherent motion, related to ventral and dorsal visual processing pathways, respectively (Braddick & Atkinson, 2011). Most of these studies used threshold detection methods to investigate visual processing. Adult levels of detection thresholds for coherent form stimuli were reached at the age of 6 years (Gunn et al., 2002), and between 6 and 9 years of age (Lewis et al., 2004). Motion detection tasks, such as Local motion and Global motion from the present study, revealed mature detection thresholds at the age of 10 years old (Gunn et al., 2002; Spencer et al., 2000). In our analysis, using RTF as outcome measure of development resulted in similar age-related effects as the mentioned studies using threshold detection methods. In addition, the current results are in line with previous findings using RTF analyses (Boot et al., 2012). Therefore, we

Stimulus	Feature channels (all weight 1)	GBVS ES %	Itti & Koch ES%
Cartoon	ORDF	8.84	5.18
Contrast	OR	8.28	5.67
Form	O	4.47	3.49
Local motion	F	4.70	2.08
Color	D	1.63	0.96
Global motion	F	0.68	0.66

Table 2. Per stimulus the effective salience (ES) values that were obtained with the GBVS and the Itti & Koch algorithm, respectively. Feature channels matched the visual content of stimuli and were all given equal weights. *Notes*: O = orientation; R = contrast; D = color; F = flicker.

conclude that using the present method to assess visually-guided orienting capacities seems a valid approach.

In stimuli with high and moderate effective salience (ES) the variation in RTF was mainly reflected in variation in RT_{min} values, i.e., the fastest speed of processing and responding to a specific visual modality. However, for stimuli with low ES and in children under the age of four years, the variation in RTF also depended on intra-individual RT variability. This variation in reflexive oculomotor responses might be ascribed to alternating visual attention in younger children.

The approach that was used to compute salience (GBVS; Harel et al., 2006) was originally designed to predict fixations on natural images and videos, which are more complex than the stimuli employed in the current study. However, this algorithm provided a best estimate of local target salience compared to its background, and thereby enabled between-stimulus comparison of salience. As a reference, Table 2 shows the ES values we obtained by computing salience values with both the GBVS and original salience algorithm (Itti & Koch, 2000). Although absolute ES values that were obtained with the original paradigm are lower, the hierarchy among stimuli was identical. Hence, our conclusion does not seem to depend on the salience algorithm. From a practical perspective, our data showed that predicting children's attention allocation may be enhanced by taking into account feature channels and developmental age in calculating salience effects.

A study limitation was the assessment of ES of the Global motion stimulus. We were not able to calculate an ES value above 1 using the GBVS standard settings described (i.e., the channel flicker). Apparently, the applied approach to calculate ES was not able to identify the target area of Global motion. Alternative salience computations are available to calculate, for example, motion-directed contrast and boundaries (Açik, Bartel, & König, 2014; Black & Anandan, 1996).

However, these methods rely on detection of multiple motions in a small image region (Black & Anandan, 1996). In our stimulus, each dot represented a local motion and together they represented an optic flow over a large area. After stimulus onset, one may notice differences in contrasts due to a relatively low dot density in the target versus the nontarget areas Indeed, when the contrast channel was activated in the salience algorithm, the ES in the target quadrant increased from 0.68 to 1.4. Possibly, salience of global motion is assessed more accurately when dot sizes or dot densities increase. Another limitation may be that Cartoons were shown 16 times, as opposed to the other stimuli that were shown four times. However, repeated presentations generally do not cause higher average RTF (Kooiker, van der Steen, & Pel, 2014), and are therefore unlikely to affect age-related development of RTF. Ideally, adult levels are used to establish the age of maturation. Certain eye movement response functions, especially those involving cortical control, are not fully mature at the age of 11 (Karatekin, 2007; Luna et al., 2008). Although we generally do not find RTF differences between our oldest age bin and a pilot group of adults, our results may need to be extended to adolescent age groups to capture the full developmental timeframe.

Conclusions

We conclude that the allocation of attention to visual targets, quantified using eye movement response times, is age-related and highly depends on visual salience of the targets. Visually-guided eye movement response times to low-salient targets improve at a slower rate and stabilize at a later age, while showing more intra-individual variability than high salient targets. These quantitative responses can be used to follow the development of visuospatial attention, and of related visual processing capacities, in children from 1 year of age.

Keywords: visual attention, salience, preferential looking, eye movements, oculomotor responses, children, development, visual information processing

Acknowledgments

Corresponding author: Marlou J. G. Kooiker.
Email: m.kooiker@erasmusmc.nl.
Address: Department of Neuroscience, Erasmus University Medical Center, Rotterdam, The Netherlands.

References

Açik, A., Bartel, A., & König, P. (2014). Real and implied motion at the center of gaze. *Journal of Vision*, *14*(1):2, 1–19, doi:10.1167/14.1.2. [PubMed] [Article]

Açik, A., Sarwary, A., Schultze-Kraft, R., Onat, S., & König, P. (2010). Developmental changes in natural viewing behavior: Bottom-up and top-down differences between children, young adults and older adults. *Frontiers in Psychology*, *1*, 207, doi:10.3389/fpsyg.2010.00207.

Atkinson, J., & Braddick, O. (2001). Gaze control: A developmental perspective. In: J. M. Zanker & J. Zeil (Eds.), *Motion vision: Computational, neural and ecological constraints* (pp. 219–225). Berlin: Springer-Verlag.

Black, M., & Anandan, P. (1996). The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, *63*, 75–104.

Boot, F. H., Pel, J. J., Evenhuis, H. M., & van der Steen, J. (2012). Quantification of visual orienting responses to coherent form and motion in typically developing children aged 0-12 years. *Investigative Ophthalmology & Visual Science*, *53*, 2708–2714. [PubMed] [Article]

Braddick, O., & Atkinson, J. (2011). Development of human visual function. *Vision Research*, *51*, 1588–1609.

Bruce, N. D., & Tsotsos, J. K. (2009). Saliency, attention, and visual search: An information theoretic approach. *Journal of Vision*, *9*(3):5, 1–24, doi:10.1167/9.3.5. [PubMed] [Article]

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron*, *58*, 306–324.

Fantz, R. L. (1965). Visual perception from birth as shown by pattern selectivity. *Annals of the New York Academy of Sciences*, *118*, 793–814.

Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: A priority map for target selection. *Trends in Cognitive Sciences*, *10*, 382–390.

Glückman, M., & Johnson, S. P. (2013). Attentional capture by social stimuli in young infants. *Frontiers in Psychology*, *4*, 527, doi:10.3389/fpsyg.2013.00527.

Gunn, A., Cory, E., Atkinson, J., Braddick, O., Wattam-Bell, J., Guzzetta, A., & Cioni, G. (2002). Dorsal and ventral stream sensitivity in normal development and hemiplegia. *NeuroReport*, *13*, 843–847.

Harel, J. (2011). *A saliency implementation in MATLAB*. Retrieved from http://www.vision.caltech.edu/~harel/share/gbvs.php

Harel, J., Koch, C., & Perona, P. (2006). Graph-based visual saliency. *Proceedings of Neural Information Processing Systems (NIPS)*. Retrieved from http://papers.nips.cc/paper/3095-graph-based-visual-saliency.pdf

Helo, A., Pannasch, S., Sirri, L., & Rämä, P. (2014). The maturation of eye movement behavior: Scene viewing characteristics in children and adults. *Vision Research*, *103*, 83–91.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506.

Kanazawa, S., Shirai, N., Ohtsuka, Y., & Yamaguchi, M. K. (2006). Perception of opposite-moving dots in 3- to 5-month-old infants. *Vision Research*, *46*, 346–356.

Karatekin, C. (2007). Eye tracking studies of normative and atypical development. *Developmental Review*, *27*, 283–348.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, *4*, 219–227.

Kooiker, M. J., Pel, J. J., & van der Steen, J. (2014). Viewing behavior and related clinical characteristics in a population of children with visual impairments in the Netherlands. *Research in Developmental Disabilities*, *35*, 1393–1401.

Kooiker, M. J., van der Steen, J., & Pel, J. J. (2014). Reliability of visual orienting response measures in children with and without visual impairments. *Journal of Neuroscience Methods*, *233*, 54–62.

Lewis, T. L., Ellemberg, D., Maurer, D., Dirks, M., Wilkinson, F., & Wilson, H. R. (2004). A window

on the normal development of sensitivity to global form in Glass patterns. *Perception*, *33*, 409–418.

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development*, *75*, 1357–1372.

Luna, B., Velanova, K., & Geier, C. F. (2008). Development of eye-movement control. *Brain & Cognition*, *68*, 293–308.

Melloni, L., van Leeuwen, S., Alink, A., & Muller, N. G. (2012). Interaction between bottom-up salience and top-down control: How salience maps are created in the human brain. *Cerebral Cortex*, *22*, 2943–2952.

Pel, J. J., Manders, J. C., & van der Steen, J. (2010). Assessment of visual orienting behaviour in young children using remote eye tracking: methodology and reliability. *Journal of Neuroscience Methods*, *189*, 252–256.

Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision*, *11*(5):9, 1–30, doi:10.1167/11.5.9. [PubMed] [Article]

Spencer, J., O'Brien, J., Riggs, K., Braddick, O., Atkinson, J., & Wattam-Bell, J. (2000). Motion processing in autism: Evidence for a dorsal stream deficiency. *NeuroReport*, *11*, 2765–2767.

Treue, S. (2003). Visual attention: The where, what, how and why of saliency. *Current Opinion in Neurobiolology*, *13*, 428–432.

Wattam-Bell, J. (1996). Visual motion processing in one-month-old infants: Preferential looking experiments. *Vision Research*, *36*, 1671–1677.

Wattam-Bell, J., Birtles, D., Nyström, P., von Hofsten, C., Rosander, K., Anker, S., . . . Braddick, O. (2010). Reorganization of global form and motion processing during human visual development. *Current Biolology*, *20*, 411–415.

World Medical Association. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *Journal of the American Medical Association*, *310*, 2191–2194.