PAPER

Blindness to background: an inbuilt bias for visual objects

Catherine G. O'Hanlon[1] and Jenny C.A. Read[2]

1. Department of Psychology, Aberystwyth University, UK
2. Institute of Neuroscience, Newcastle University, UK

Abstract

Sixty-eight 2- to 12-year-olds and 30 adults were shown colorful displays on a touchscreen monitor and trained to point to the location of a named color. Participants located targets near-perfectly when presented with four abutting colored patches. When presented with three colored patches on a colored background, toddlers failed to locate targets in the background. Eye tracking demonstrated that the effect was partially mediated by a tendency not to fixate the background. However, the effect was abolished when the targets were named as nouns, whilst the change to nouns had little impact on eye movement patterns. Our results imply a powerful, inbuilt tendency to attend to objects, which may slow the development of color concepts and acquisition of color words.

Research highlights

- Toddlers show a striking deficit in their ability to point to target colors when the color in question is in the background of an image.
- Two- to 4-year-olds were able to point accurately to named colors when these were presented as foreground patches, but not when the same colors formed the background of a visual display, despite the fact that the images' background occupied at least two-thirds of the display. This effect was much less strong in older children aged 5 to 12 years and in adults.
- The 'blindness to background' effect was abolished by the use of nouns, for example when asking children to point to 'water' rather than to 'blue', whilst eye movement patterns did not differ across adjective and noun conditions. This shows that the inability to switch attention from figure to ground depends on the nature of the linguistic cue, providing robust, converging evidence of a powerful, inbuilt tendency to attend to objects when interpreting color language.
- Eye tracking was achieved with children as young as 2 years.

Introduction

From a very young age, children acquire vocabulary in a spontaneous and intrinsically motivated fashion (Golinkoff, Hirsh-Pasek, Bloom, Woodward, Akhtar *et al*., 2000; Woodward & Markman, 1998). This has been shown to be true of nouns, adjectives and verbs well before a child's second birthday (Waxman, 2002). However, an atypical degree of effort has been documented in the learning of color adjectives, and therefore in the development of color concepts (Bornstein, 1985; Kowalski & Zimiles, 2006; Rice, 1980; Soja, 1994). The current study aimed to provide complementary evidence to explain children's slow and effortful development of color concepts and acquisition of color vocabulary. The focus is on attentional processes underpinning lexical acquisition, which are central to word–object mapping. The study was motivated by three lines of research in this domain.

First, from the onset of productive vocabulary, children are quick at figuring out word meanings ('fast-mapping'; Carey, 1978; Markman, 1989), particularly that of count nouns. A typical 2-year-old will learn the meaning of a noun following a single exposure to it in a

Address for correspondence: Catherine O'Hanlon, Department of Psychology, Aberystwyth University, Penbryn 5, Penglais Campus, Aberystwyth, SY23 3UX, UK; e-mail: cao15@aber.ac.uk or Jenny Read, Institute of Neuroscience, Henry Wellcome Building, Framlington Place, Newcastle University, Newcastle Upon Tyne, NE2 4HH, UK; e-mail: jenny.read@newcastle.ac.uk

context of ostensive definition (e.g. 'Look, a dog!'), guided by the 'whole object bias' (Woodward & Markman, 1998). In addition, extension errors, such as using the word 'dog' to name other four-legged animals (e.g. McDonough, 2002), are overcome with minimal training or feedback.

Second, whilst words that describe object properties are generally harder to learn, appear later than count nouns in children's vocabularies (e.g. Fenson, Dale, Reznick, Bates, Thal *et al.*, 1994) and are less frequent in parental language during the early years (Bergelson & Swingley, 2013), there are differences in the rates of learning across different domains. For example, shape words have been shown to be much easier to learn compared to color words, when taught within the same learning environment (O'Hanlon & Roberson, 2006, 2007), and several studies demonstrated a tendency to extend novel words on the basis of shape when color, shape, texture or size were viable interpretations of a new word (e.g. Au & Markman, 1987; Baldwin, 1989; Landau, Smith & Jones, 1988). This has been said to show selective attention to shape in word learning (Gershkoff-Stowe & Smith, 2004; Landau *et al.*, 1988; Landau, Smith & Jones, 1992a). In addition, 2-year-olds' ability to match objects on the basis of color is disrupted when the dimensions of shape or texture vary (Soja, 1994). This is presumably at least in part because toddlers' attention is grasped by object shape or object texture whilst color loses salience. Moreover, infants' learning of object categorization requires more training when based on surface structure than it does when based on object shape (Gliga, Mareschal & Johnson, 2008).

Third, children acquire vocabulary at the same time or *after* developing the corresponding underlying concepts (Mandler, 2004). This means that the brain wires up to associate perceptual, semantic and verbal information that is then correctly generalized and applied beyond the context and specific objects in view at the time of learning. However, contrary to what we see in children's acquisition of count nouns and dimensional adjectives in domains other than color, color words appear to be learned before the underlying concepts are developed (Kowalski & Zimiles, 2006; Sandhofer & Smith, 1999), and extension errors are overcome only after training and reinforcement are given over a prolonged period of time (e.g. Rice, 1980; Roberson, Davidoff, Davies & Shapiro, 2004).

Developmental researchers have attempted to explain the reasons underpinning the apparently atypical developmental pattern in children's learning of the color domain using a variety of tasks, manipulating the objects used (Kowalski & Zimiles, 2006; Pitchford & Mullen, 2001; Sandhofer & Smith, 1999; Soja, 1994) or the

(verbal) instructions given (Baldwin, 1989; Borstein, 1985; Gottfried & Tonks, 1996; O'Hanlon & Roberson, 2006; Rice, 1980; Waxman & Markow, 1998). Non-verbal tasks have typically involved matching or sorting objects by color; verbal tasks have primarily attempted to teach new color words. The most striking example of young children's difficulties at learning the meaning of color words comes from studies that have used linguistic contrast (e.g. Au & Laframboise, 1990; Gottfried & Tonks, 1996; O'Hanlon & Roberson, 2006). In this paradigm, a 3-year-old might be told, 'This is mauve; it's not blue', where 'mauve' is the word to be learned and 'blue' is a known word (firm knowledge of the contrasted color is established in pretests). Next, the researcher may point to an array of, say, six objects, two matching the first on texture, two matching it on shape, and two on color, and asks the child, 'Can you show me another mauve one?' Regardless of the semantic cue given in the naming event ('This is X; it's not *blue*'), the child is highly likely to base the response on shape similarity. This behavior would seem to suggest that the child ignored the information provided to guide learning, but children's attention is immediately drawn to verbal information offered within socio-communicative contexts, even when speech is not directly addressed to them (e.g. Akhtar, 2005; Akhtar, Carpenter & Tomasello, 1996). Toddlers can use color information implicitly even before they demonstrate understanding of color words. Johnson, McQueen and Huettig (2011) found that words which are associated with a prototypical color caused 2-year-olds to make eye movements to distractor objects with that color, while naming the color did not. For example, toddlers would look more often at a yellow cup than a red cup when asked 'Can you find the banana?', even though they showed no significant difference when asked 'Where is the yellow one?'

Together, these results suggest that young children have a strong tendency to attend to whole objects in their environment. In identifying objects, children give most weight to shape information and less weight to color information. Children find it particularly difficult to conceptualize color independently of objects; e.g. using 'yellow' to guide visual search when primed with the noun 'banana' but not with the adjective 'yellow' (Johnson *et al.*, 2011).

In the current study, we investigated the possibility that a strong bias to attend to discrete entities in one's environment might result in inattention to color even when the linguistic context should direct attention to the latter alone. We implemented a novel visual search task and tracked eye movement patterns during search times. Children as young as 2 years of age viewed colorful visual images on a touch-sensitive computer monitor;

they were trained to touch the screen to indicate where they saw a named color or object, following a one-word instruction. This was either a color adjective (e.g. 'Green!' Experiment 1) or a noun (e.g. 'Grass!' Experiment 2). Eye movement patterns were recorded remotely as participants scanned the image prior to their pointing response. In most trials, the named target was present in the image; in one-third of trials the target was absent. Children were trained to touch a 'sad face' also presented on the touchscreen monitor if they believed the named target was not present. Thus, two critical manipulations were implemented. First, the target, when present, was either a foreground object or the image's background. Second, attention to targets was elicited either by a color word or a noun. Critically, children were pre-tested to ascertain that they had firm knowledge of color targets. This enabled us to address the hypothesis that early color concepts are intrinsically associated to discrete objects. In line with this, when cued with a color word, 2- to 4-year-olds were expected to struggle to locate target colors when present in the background, but not when present as figure. Older children (5 to 12 years) and adults were expected to be accurate on all trials (foreground, background, absent), as the displays were visually rather simplistic, suitable for children as young as 2. When cued with a noun, all children were expected to locate foreground and background targets with equal ease.

In terms of eye movement patterns, and particularly with the younger children, we anticipated one of two possible outcomes. In line with the literature on inattentional blindness in both adults (Simons, 2000) and children (Memmert, 2006; Remington, Cartwright-Finch & Lavie, 2014), fixations to the background of visual displays might be evident even when participants might report (present) background targets as being absent. Alternatively, a tendency to attend to foreground objects might result in fewer fixations to the stimuli's background, particularly when participants were cued with color words.

General methods

Participants

All participants were white, native English-speaking Caucasians from a middle-class socioeconomic background and none had a significant sensory impairment. Child participants were recruited by distribution of study information packs at nurseries and primary schools in Newcastle upon Tyne. Adult participants were recruited from within the University environment. We analyzed

results in four different age groups: 2- to 3-year-olds; 4-year-olds; 5- to 12-year-olds; and adults. Details of the participants in each group who did each experiment are given in Table 1.

Experiments

All stimuli consisted of four colors: one target color and three other colors.

Control task: Patchwork display with adjective cues

The Control task was played first; it also served as a training session to teach children how to play the games. The visual stimuli consisted of a 'patchwork' of four different colors (Figure 1A). Participants then went on

Table 1 *Participant statistics. Participants first did the Control task and then, ideally, completed Experiment 1 and either the Adjective or Noun condition of Experiment 2, in random order. Not all children were willing to complete all 4 tasks, which is why N for Experiment 1, or the total N for the two conditions of Experiment 2, are sometimes lower than N for Control*

| Age group	Number of participants	Number males + females	Mean age (years)	SD age (years)
2–3-year-olds				
Control task	26	14M + 12F	3.19	0.51
Experiment 1: Squares	20	10M + 10F	3.10	0.51
Experiment 2: Adjectives	12	7M + 5F	3.09	0.40
Experiment 2: Nouns	13	6M + 7F	3.22	0.59
4-year-olds				
Control task	18	7M + 11F	4.47	0.28
Experiment 1: Squares	16	6M + 10F	4.50	0.27
Experiment 2: Adjectives	10	3M + 7F	4.53	0.31
Experiment 2: Nouns	8	4M + 4F	4.40	0.23
5–12-year-olds				
Control task	25	18M + 7F	7.82	2.08
Experiment 1: Squares	22	15M + 7F	8.19	1.93
Experiment 2: Adjectives	12	8M + 4F	7.83	2.30
Experiment 2: Nouns	13	10M + 3F	7.81	1.96
Adults				
Control task	30	6M + 24F	21.03	1.48
Experiment 1: Squares	30	6M + 24F	21.03	1.48
Experiment 2: Adjectives	15	5M + 10F	20.54	1.29
Experiment 2: Nouns	15	1M+14F	21.52	1.53

(a)

(b)
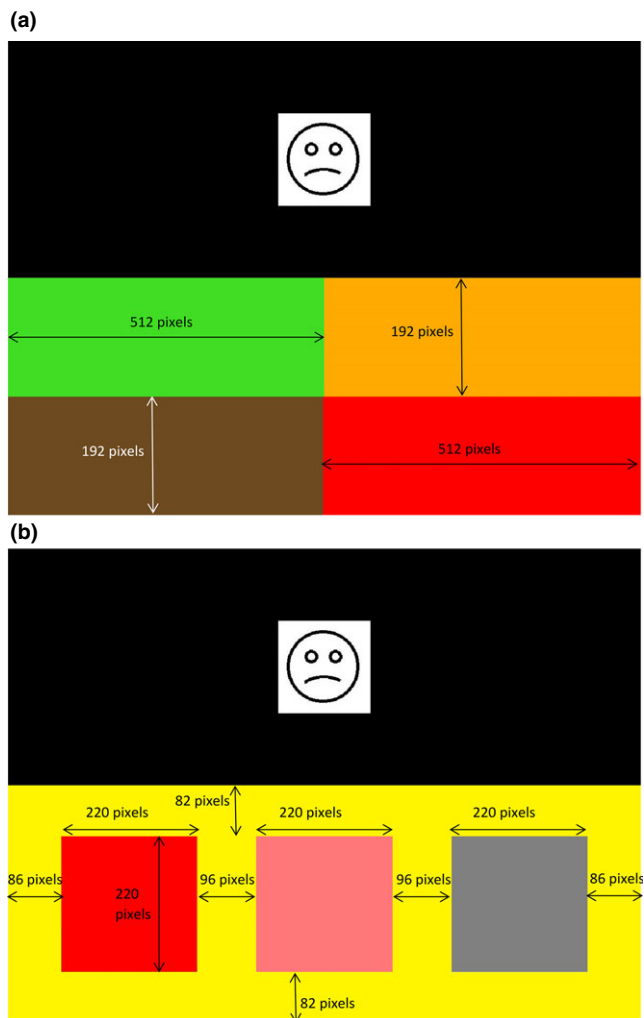
Figure 1 *Panel A: Example visual displays of a Control Patchwork task trial. Clockwise from the top left, the colors are green, orange, red and brown. The target color in this display must be green (brown and orange were never targets, and orange was not presented when red was a target). Panel B: Example visual display during a trial from the Squares task. The background is yellow, and the foreground objects from left to right are red, pink and gray. The target color in this display must be yellow (pink and gray were never targets, and pink was not presented when red was a target). The arrows and dimensions are given here for reference only; they were not present in the displays.*

to the two experiments described below, in a random order. That is, all participants who were willing to cooperate completed Experiment 1 (Squares task), and Experiment 2 (Scenes task, with either Adjective or Noun cues). See Table 1 for the numbers completing each experiment.

Experiment 1: Squares task with adjective cues

The visual stimuli consisted of three squares (the foreground objects) laid out on a uniform-color background (Figure 1B). Different colors were used for each square and the background.

Experiment 2: Scenes task with adjective and noun cues

The visual stimuli depicted meaningful scenes (house, bus, playground, aquarium), which parents confirmed were familiar to our participants. The 12 images used are shown in Table 2. Experiment 2 had two conditions (Adjective, Noun); each child was randomly assigned to one of these. In the Adjective condition, the target was indicated by a recorded voice naming a color (e.g. 'Blue!'). In the Noun condition, the target was indicated by a recorded voice saying a noun (e.g. 'Water!'). The 12 nouns used are given below each image in Table 2.
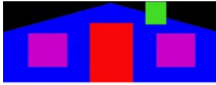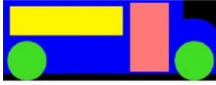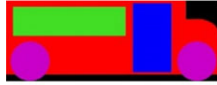
Apparatus and materials

Experiments were run on a Dell PC, using Matlab (www.mathworks.com) with the Psychophysics Toolbox extensions (Kleiner, Brainard, Pelli, Ingling, Murray et al., 2007; Pelli, 1997) and a modified version of the iViewX toolbox developed by Frans Cornelissen. Visual stimuli were presented on a touch-sensitive 15-inch LCD monitor (mod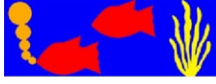el CVJU-E38 from Chinavision) with a resolution of 1024 × 768 pixels. Eye movements were monitored using a 50 Hz remote eye tracker (RED) from SensoMotoric Instruments (http://www.smivision.com/en/gaze-and-eye-tracking-systems/products/red-red250-red-500.html). Children rested their head against a pressure-sensitive button switch (Ablenet, http://www.ablenetinc.com/Assistive-Technology/Switches/Big-Red). Logitech speakers output target words. Stimuli used best examples (Berlin & Kay, 1969) of red, blue, green, yellow, pink, purple, orange, brown and gray, generated in Matlab (Experiment 1) and in Photoshop CS2 (Experiment 2).

Design

On each trial, the participant was shown the visual stimulus on a touch-sensitive monitor and simultaneously given a target via a one-word auditory instruction. The participant had to search for the target in the visual stimulus, touch it if present, or touch a target-absent icon (a 'sad face') if absent. The experiments were run by computer. The computer selected the order of trials, chose the colors to present, displayed the stimuli and played the audio instruction over speakers. The

Table 2 *Scenes used in Experiment 2. The words under each image are the spoken instruction given in the Noun condition. In the Adjective condition, the spoken instruction was the name of the target color, e.g. 'red' for all 3 scenes in the top row. The 12 images were presented in the bottom half of the touchscreen monitor; the top half of the monitor displayed a 'sad face' as shown in Figure 1*

| Color | Target Location | | |
	Foreground	Background	Absent
"Red"	"Door"	"House"	"Cat"
"Yellow"	"Window"	"Bus"	"Dog"
"Green"	"Tree"	"Grass"	"Swing"
"Blue"	"Bubbles"	"Water"	"Spider"

instruction was repeated at 3-second intervals until the child responded. The computer recorded the location where the child touched the screen, and the reaction time from the offset of the initial instruction to the touch.

There were four possible target colors: red, blue, green and yellow. Pink, purple, orange, brown and gray were filler colors, never used as targets. To avoid task ambiguity, we applied the following constraints: (i) if the target was red, the stimulus could not contain orange or pink; (ii) if the target was blue, the stimulus could not contain purple or green; (iii) if the target was yellow, the stimulus could not contain orange; (iv) if the target was green, the stimulus could not contain blue. This demonstrably avoided ambiguity, since all age groups achieved near 100% accuracy in the Control task.

In the two experiments, there were three types of trials: target-absent trials, in which the named target color was not present; foreground-present trials, in which it was present as a foreground object; and background-present trials, in which it was the background. This resulted in a total of 3(target present in foreground, target present in background, target-absent) × 4(red, blue, green, yellow) = 12 trials per experimental task. The Control task contained no distinction between foreground and background (Figure 1A), so the 12 trials consisted of four target-absent and eight target-present trials.

The computer displayed the 12 trials in a random order subject to four constraints: (i) consecutive trials never had the same target color; (ii) consecutive trials never had the same target location; (iii) there were never more than two consecutive trials on which the target was absent; (iv) on the experimental tasks, there were never more than three consecutive trials on which the target was not present in the foreground (i.e. background or absent).

The experiments were blocked. Participants first completed the Control task, and then either played the full set of 12 trials for Experiment 1 followed by the full set of 12 trials for Experiment 2, or vice versa, in line with the counterbalancing of task order. Experiment 1 used a within-groups design (Control, Squares task with adjective cues). Experiment 2 used a between-groups design (Scenes task: Noun and Adjective cues). From a statistical point of view, it would have been desirable to use a within-subjects design for both experiments and to include more than four trials per subject per condition, but we were limited by the abilities and interest of our very young participants.

Procedure

The study took place in a child-friendly laboratory environment. Child participants completed pretests, control and experimental tasks on two separate occasions, exactly one week apart. During the first visit, written parental consent was obtained and children's color knowledge was tested using naming ('What color is this?') and comprehension ('Can you show me the blue one?') tasks using hand-held objects. Participation in Experiments 1 and 2 was subject to correctly naming and picking all four target colors (red, blue, green, yellow). On the second visit, children completed up to three computerized tasks (Control, Experiment 1, Experiment 2) in a window-less room lit only by a dim floor lamp. A research assistant, naïve to the study aims and unable to see the experimental stimuli, was present throughout to maintain shared interest in the task with the child. If a child hesitated in giving a response, the researcher said, 'If you can see X touch it; if you cannot see X touch the sad face', where X was the one-word instruction played by the computer.

Children sat on a car seat fixed to a chair; they faced the touchscreen monitor located at a distance of about 50 cm from the child's forehead. Speakers were located to the left and to the right of the monitor, respectively. The eye tracker (RED) was located below the monitor, looking up at the child from a distance of about 65 cm. The stimuli were presented only in the bottom half of the monitor (see Figure 1), since the lower eyelids blocked the eye tracker's view of the pupil when the children looked at the upper half. The 'sad face' was displayed in the top half of the monitor.

Successful eye tracking required children to keep their head within a narrow window and not to block the tracker's view with their hands. In between trials and while viewing the stimuli, children were trained to sit with their hands on their knees and their head resting on a round button switch (12.7 cm diameter) interfaced to the computer. A trial would not start until the child's head rested on the button switch. Most of the younger children, having short arms, had to lean forward to give their response, at which point the tracker lost sight of their eyes. After the response was given, a beep sound started and continued until the child returned to the correct position in the car seat.

Eye tracking

Calibration

The eye tracker was calibrated for each child immediately before each task. An animated clown face popped up at five locations on the bottom half of the touchscreen monitor, where the experimental stimuli were displayed. During calibration, the computer played a recording of a child's voice calling. In most cases, we also pre-recorded the individual child's name and mixed that in, so the child might hear, for example 'Michael! Look over here! Michael! Hello! Can you see me? <laughter> Michael!' These techniques generally succeeded in getting the child to fixate the calibration target as required.

Analysis

Eye tracking was technically challenging and eye position data were not successfully recorded for all children nor at every point throughout the experiment. When the RED lost the signal, it tended to report eye positions close to the origin of its coordinate system, corresponding to the top-left corner of the stimuli. To avoid mistakenly analyzing these as real eye positions, in our analysis we discarded data points within 20 pixels of the origin. Fixations off the screen, to the top half of the screen, or to the black regions in the 'bus' and 'house' displays (see Table 2) were discarded from the analysis. For simplicity, we computed the proportion of all valid eye positions recorded by the eye tracker (i.e. time points) for each region of interest for each participant. We will refer to all eye positions as fixations, although our analysis did not distinguish between whether a recorded eye position was part of a fixation or saccade. Since the RED recorded data at 50 Hz, we do not have the temporal resolution to capture saccades. We classified fixations as being to the background or foreground and, where appropriate, whether they were to the particular foreground object designated as the target or to another foreground object. To do this, we defined foreground regions of interest that were about 20 pixels larger than the foreground objects. This is because viewers often scan along the edge of an object (Henderson, 2003).

Statistical analysis

For both Experiments 1 and 2, mixed Analysis of Variance (ANOVA) was used to examine overall differences in accuracy and reaction time (separately), with age group (2–3-year-olds, 4-year-olds, 5–12-year-olds, adults) entered as the between-groups factor, and trial type (foreground, background, absent) entered as the within-groups factor. Experiment 2 also included condition (Noun, Adjective) as a between-groups factor. Significant main effects of age were followed up with Tukey HSD; significant interactions and within age

group differences were analyzed using paired samples *t*-tests, applying the Bonferroni adjustment for multiple comparisons where appropriate. This is indicated where necessary.

Experiment 1

Two- to 12-year-olds and adults completed computerized tasks adopting visual stimuli displaying four colors either as abutting patches (control task) or as three squares over a uniform-color background (experimental task). They were required to locate a target that was either figure or ground, specified by a one-word instruction naming the target's color (e.g. 'Green!'). The aim was to investigate whether toddlers find it hard to locate a known color when this is presented outside the context of an object, and whether they differ in this respect from older children and adults. Eye tracking was used to examine attentional patterns during visual search.

Results and discussion

Accuracy

Figure 2 summarizes accuracy scores on the Control Patchwork and the experimental Squares tasks, across age groups. For the Control task, accuracy and reaction times (RTs) were averaged across the four target colors for target-present and target-absent trials. These data are provided in Table 3. Similarly, for the Squares task, accuracy and RTs were averaged across the four targets for foreground, background and absent trials, respectively; data are given in Table 4.

The left-hand-side panels of Figure 2 show accuracy data on the Control Patchwork task. These ranged from 92% correct in the youngest age group to 98% correct among the older children and adults. Paired samples *t*-tests (p = [.05/4] .013 for statistical significance) showed no significant differences between target-present and target-absent trials among any age group (2- to 3-year-olds: $t(25) = 0.960$, $p = .346$; 4-year-olds: $t(17) = 0.622$, $p = .542$; 5- to 12-year-olds: $t(24) = 1.809$, $p = .083$; adults, $t(29) = 2.112$, $p = .043$). These data confirm knowledge of target colors, understanding of the task, and being able to report the target as absent when it was absent, particularly for the youngest age group.

The right-hand-side panels of Figure 2 display accuracy data on the Squares task. Here, strong differences emerged between trials on which the target was present in the background compared to trials on which it was present in the foreground, or was absent. A mixed-design Analysis of Variance computed on the Squares task with one 4-level between-subjects factor (2–3-year-olds, 4-year-olds, 5–12-year-olds, adults) and one 3-level within-subject factor (target present in foreground, target present in background, target absent) revealed significant main effects of age group, $F(3, 84) = 15.289$, $p < .001$, and trial type, $F(2, 168) = 85.825$, $p < .001$, and a significant interaction between age group and trial type, $F(6, 168) = 5.287$, $p < .001$. Tukey HSD showed that, overall, the 2- to 3-year-olds performed significantly worse than the 4-year-olds ($p = .028$), the 5- to 12-year-olds and the adults (both $p < .001$). Four-year-olds' performance differed from that of the 5- to 12-year-olds ($p = .028$) but not from that of the adults ($p = .136$). Adults' performance was also similar to that of the 5- to 12-year-olds ($p = .794$). Within-group comparisons (see Figure 2, panels A, B, C and D, respectively) of accuracy on foreground vs. background trials (p = [.05/4] .013 for statistical significance) showed that the 2- to 3-year-olds performed significantly worse on background trials (21% correct), $t(19) = 7.740$, $p < .001$, as did the 4-year-olds (39% correct), $t(15) = 5.115$, $p < .001$, and the adults (67% correct), $t(29) = 4.042$, $p < .001$, but not the 5- to 12-year-olds (81% correct), $t(21) = 2.306$, $p = .031$. Thus, although adult error rates were higher on background trials relative to those seen among the older children, the difference was not statistically significant, as shown by an independent samples *t*-test, $t(50) = 1.167$, $p = .249$.

No within age group differences were found in accuracy on foreground trials (ranging from 88% to 100% correct) vs. absent trials (ranging from 84% to 100% correct; all $p > .05$). Instead, comparisons between absent and background trials showed significantly worse performance on the latter among the 2- to 3-year-olds, $t(19) = 7.611$, $p < .001$, the 4-year-olds, $t(15) = 5.115$, $p < .001$, and the adults, $t(29) = 4.042$, $p < .001$, but not among the 5- to 12-year-olds, $t(21) = 1.945$, $p = .065$.

For the subjects who performed both tasks, we were able to directly compare performance on the Control Patchwork and Squares tasks. Eight paired samples *t*-tests (p = [.05/8] .006 for statistical significance) were used to compare accuracy on the Control Patchwork-present trials vs. the Squares foreground trials, and between Patchwork-present vs. Squares background trials, within age groups. The 2- to 3-year-olds performed significantly worse on Squares background relative to Patchwork-present trials, $t(19) = 9.099$, $p < .001$, as did the 4-year-olds, $t(15) = 4.999$, $p < .001$, and the adults, $t(29) = 4.042$, $p < .001$, but not the 5- to 12-year-olds, $t(21) = 2.306$, $p = .031$. No differences were found between Patchwork-present and Squares foreground trials among any of the four age groups (all $p > .05$). Thus, performance is equally good whether target colors are presented in an arrangement with no obvious figure/
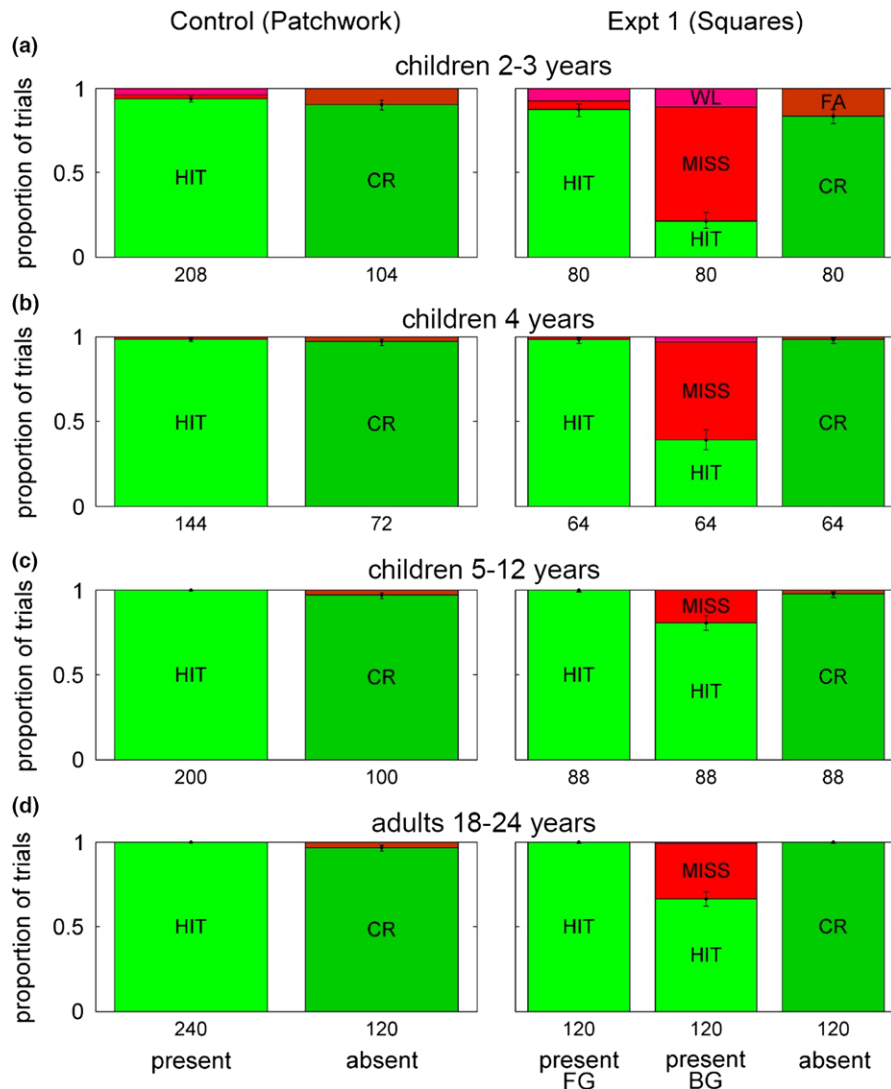
Figure 2 *Accuracy on the Control Patchwork task (left) and in Experiment 1 (Squares task, right). The two bars on the Patchwork panels show results for the eight 'target present' trials and the four 'target absent' trials completed by each subject, respectively. The three bars on the Squares panels show results for the four 'target present in the foreground' trials (present FG), the four 'target present in the background' trials (present BG), and the four 'target absent' trials. The four rows are for different age groups, as labeled. Green shades = proportion of trials answered correctly (hits or correct rejects, CR); red shades = proportion of trials answered incorrectly (misses, false alarms, FA, or wrong location, WL). The type of trial is marked on the bar where space permits. The numbers below each bar show the total number of trials done in that condition. The total number of subjects can be found by summing the numbers under all bars and dividing by 12. For child participants, this is slightly higher in the Patchwork task as not all of them proceeded to the Squares task. Error bars show the 68% confidence interval on the probability of answering correctly, assuming simple binomial statistics.*

ground structure, as in the Patchwork, or as discrete objects. Rather, participants have difficulties specifically when the target color is presented as the background.

In sum, statistical analyses computed on the accuracy data of Experiment 1 demonstrate that our two youngest age groups, and to a lesser extent the older participants, found it hard to locate a target color when this was

presented in the background (i.e. outside the context of an object).

Reaction time (RT)

Figure 3 shows the reaction times on the Patchwork and Squares tasks, plotted separately for trials on which

Table 3 *Results of the Control experiment. Reaction times are in seconds; standard error is given in parentheses. Results are given for all subjects tested on the Control task (values plotted in Figures 2 and 3), and also, for comparison, for the subset of subjects who completed Experiment 1 in addition to the Control task. There are no significant differences when we restrict ourselves to this subset*

Age group	All subjects tested on Control: Patchwork task			Only subjects who also completed Experiment 1: Squares task		
	N subjects	Target present	Target absent	N subjects	Target present	Target absent
2–3-year-olds						
Number of trials:	26	208	104	20	160	80
Proportion of trials correct		0.94	0.90		0.93	0.89
Reaction time on correct trials		3.26 (0.13)	5.58 (0.24)		3.41 (0.16)	5.75 (0.29)
Reaction time on incorrect trials		3.99 (0.31)	5.07 (0.48)		4.26 (0.29)	5.29 (0.48)
4-year-olds						
Number of trials:	18	144	72	16	128	32
Proportion of trials correct		0.99	0.97		0.99	0.97
Reaction time on correct trials		2.80 (0.33)	3.82 (0.33)		2.80 (0.36)	3.92 (0.36)
Reaction time on incorrect trials		4.07 (0.77)	5.74 (1.42)		5.15 (-)	5.74 (1.42)
5–12-year-olds						
Number of trials:	25	200	100	22	176	88
Proportion of trials correct		1.00	0.97		1.00	0.97
Reaction time on correct trials		2.00 (0.07)	2.56 (0.11)		1.93 (0.07)	2.42 (0.09)
Reaction time on incorrect trials		–	3.65 (0.78)		–	3.65 (0.78)
Adults						
Number of trials:	30	240	120	30	240	120
Proportion of trials correct		1.00	0.97		1.00	0.97
Reaction time on correct trials		1.45 (0.05)	1.71 (0.04)		1.45 (0.05)	1.71 (0.04)
Reaction time on incorrect trials		–	1.61 (0.03)		–	1.61 (0.03)

participants answered correctly (green bars) and incorrectly (red bars).

For the subjects who performed both tasks, eight paired samples *t*-tests (p = [.05/8] 0.006) were used to compare correct-trial reaction times within age groups. No differences were found between reaction times in the Patchwork and Squares tasks in any age group, regardless of whether we compared Patchwork-present vs. Squares foreground (all $p > .01$), Patchwork-present vs. Squares background trials (all $p > .09$) or Patchwork-absent vs. Squares absent (all $p > .03$). Note that for Squares background trials, the power to detect a difference is reduced because 34/88 participants did not answer any of these trials correctly. This again indicates that the arrangement of colors into Squares as opposed to Patchwork had no particular effect on participants except for their tendency to miss target colors in the background.

Considering the Squares task only, Analysis of Variance for the 54 subjects who performed at least one accurate trial in all conditions showed that, overall, reaction times on accurate trials differed significantly across age groups, $F(3, 50) = 53.885$, $p < .001$, and trial types, $F(2, 100) = 35.753$, $p < .001$. The interaction between age group and trial type was statistically significant, $F(6, 100) = 6.378$, $p < .001$. Tukey HSD showed significant differences between all age groups, with *p*-values ranging from $p = .021$ to $p < .001$.

Unsurprisingly, irrespective of age, participants took longer to answer when the target was absent (foreground 'hit' [Mean RT = 1.96 s] vs. absent 'correct reject' [Mean RT = 2.89 s], $t(87) = 9.044$, $p < .001$, paired *t*-test), presumably because they spent additional time scanning the image in search of the target. What is interesting is that participants were also significantly slower when the target was present but in the background, even on trials where they did correctly identify it in the end (foreground 'hit' [Mean RT = 1.894 s] vs. background 'hit' [Mean RT = 2.494 s], $t(53) = 4.396$, $p < .001$; this paired *t*-test is only for the 54/88 participants who answered at least one background trial correctly). This difference was most substantial and statistically significant in the youngest age group, (foreground 'hit' [Mean RT = 3.334 s] vs. background 'hit' [Mean RT = 5.107 s], $t(7) = 3.808$, $p = .007$, paired *t*-test), but persisted even in the adults (foreground 'hit' [Mean RT = 1.345 s] vs. background 'hit' [Mean RT = 1.528 s], $t(20) = 2.497$, $p = .021$, paired *t*-test), despite the fact that the background was far greater in area than that cumulatively occupied by the three squares, so one might have expected reaction times to be faster on background trials among this group.

On trials where the target was present in the background, there was little difference in reaction time between trials on which participants correctly touched the target versus trials on which they incorrectly touched

the sad face (background 'hit' [Mean RT = 3.923 s] vs. background 'miss' [Mean RT = 3.356 s], $t(12) = 1.371$, $p = .195$, paired t-test for the 13 participants who had at least one hit and at least one miss on background trials). Similarly, for trials on which participants touched the sad face to indicate an absent target, there was little difference in reaction time between trials on which the target was not present versus trials on which it was present in the background (absent 'correct reject' [Mean RT = 3.372 s] vs. background 'miss' [Mean RT = 3.280 s]), $t(546) = 0.629$, $p = .533$, paired t-test). This suggests that the high error rate in the background condition was not just because children (and adults) responded too quickly. They appear to have been making a real effort to locate target colors.

Overall, these data suggest a tendency to favor figure over ground in a visual search where attention is directed to color. This tendency is particularly striking in toddlers, especially given that they often scan the image for several seconds and still fail to locate a target present in the background. However, the accuracy data indicate that it persists to some extent in older children and even more so in adults. In particular, adults' error rates on background trials were not dissimilar to those seen for the 4-year-olds.

Eye tracking

For many subjects, eye movements were successfully recorded as they performed the task. Figure 4 shows the average proportion of fixations to different areas of the image. Fixations are classified as being to foreground objects (colored bars in Figure 4) or to the background (gray bars). For trials where the target was a foreground object, fixations are further classified as being to the target object (dark blue bars) or to another foreground object (cyan). The section of each bar corresponding to fixations onto the target, where present, is outlined in red. Analysis of Variance found a significant effect of age group, $F(3, 205) = 8.246$, $p < .001$, and trial type, $F(2, 205) = 16.62$, $p < .001$; the interaction between age group and trial type was not significant, $F(6, 205) = 0.073$, $p = .998$ (two-way ANOVA on the proportion of fixations to background on correctly-answered trials only). Tukey HSD showed significant differences between all target locations, with p-values ranging from $p = .013$ to $p < .001$, but a significant difference only between the adults and each of the three child age groups ($p < .004$ in all cases). Comparing adults and the youngest participants, we see that adults directed a larger proportion of fixations to the background than the two youngest age groups.

As one might have anticipated, when the target was present, a higher proportion of fixations were directed towards it on trials where the target was successfully identified than on those on which it was missed (the bars outlined in red in Figure 4 are always longer on the 'correct' than on the 'wrong' trials, and the confidence intervals show this is highly significant ($p < .001$)). This suggests that many misses were due to the participant not fixating the background.

Figure 5 shows eye tracking data for a 2-year-old who, unusually, correctly identified background targets on 3 out of the 4 trials in which this occurred. It is clear that

Table 4 *Results of Experiment 1: Squares. Reaction times are in seconds; standard error is given in parentheses*

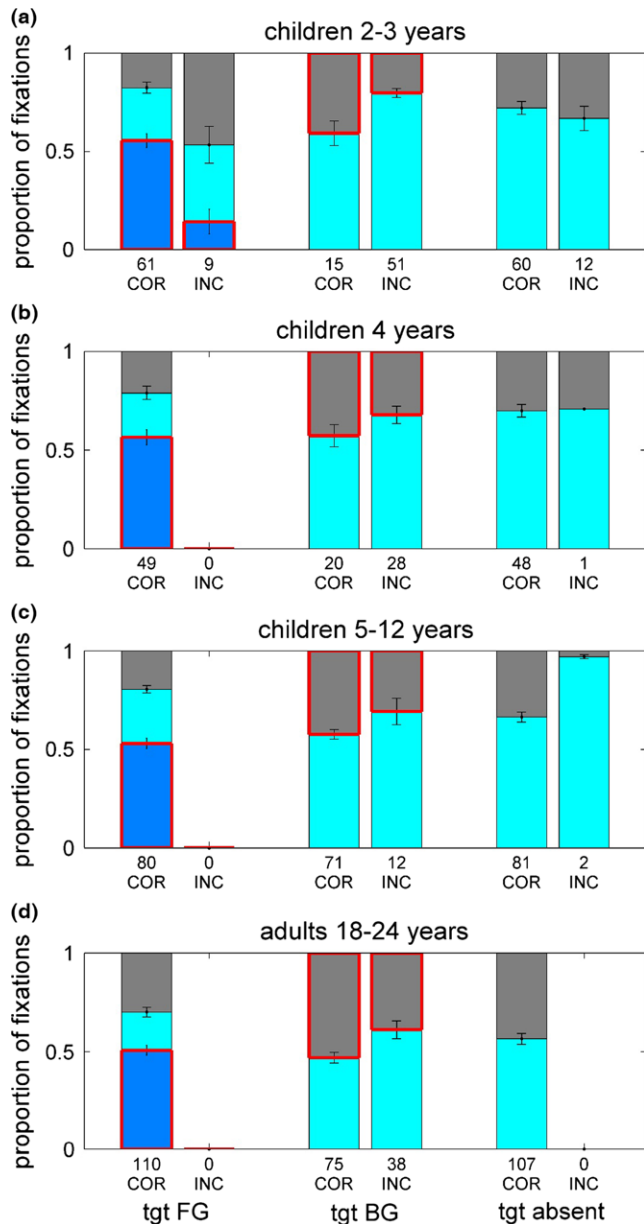| Age group	N subjects	Target present in foreground	Target present in background	Target absent
2–3-year-olds				
Number of trials:	20	80	80	80
Proportion of trials correct		0.88	0.22	0.84
Reaction time on correct trials		2.92 (0.15)	4.97 (0.44)	5.05 (0.29)
Reaction time on incorrect trials		4.07 (0.48)	4.58 (0.19)	4.00 (0.59)
4-year-olds				
Number of trials:	16	64	64	64
Proportion of trials correct		0.98	0.39	0.98
Reaction time on correct trials		2.19 (0.08)	3.17 (0.56)	3.40 (0.18)
Reaction time on incorrect trials		4.06 (–)	3.53 (0.26)	6.94 (–)
5–12-year-olds				
Number of trials:	22	88	88	88
Proportion of trials correct		1.00	0.81	0.98
Reaction time on correct trials		1.79 (0.06)	2.15 (0.16)	2.38 (0.08)
Reaction time on incorrect trials		–	2.34 (0.19)	2.14 (0.16)
Adults				
Number of trials:	30	120	120	120
Proportion of trials correct		1.00	0.67	1.00
Reaction time on correct trials		1.32 (0.02)	1.51 (0.06)	1.58 (0.04)
Reaction time on incorrect trials		–	1.61 (0.08)	–
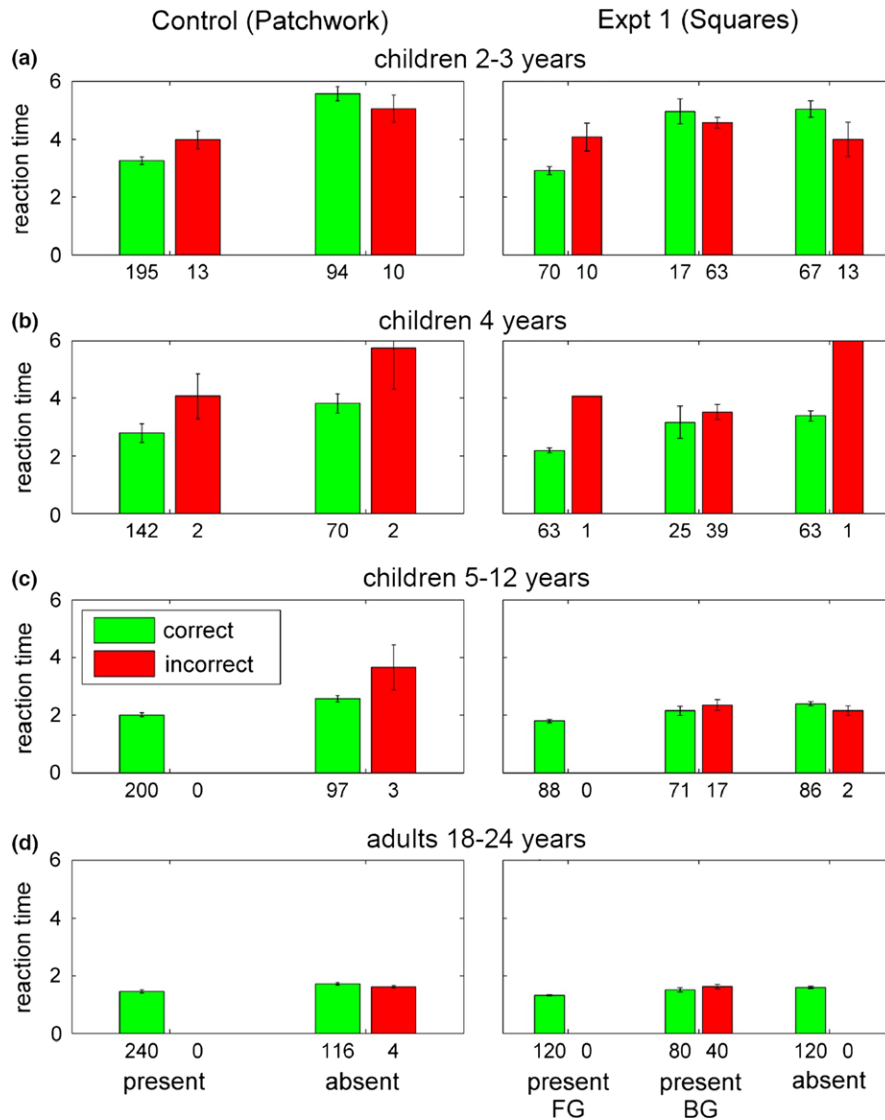
Figure 3 *Reaction times (seconds) on the Control Patchwork task and the Squares task of Experiment 1. Green bars show mean reaction time on trials where participants answered correctly; red bars show mean reaction time on trials where participants answered incorrectly (if any). The numbers below each bar show the total number of trials done in that condition. The total number of subjects can be found by summing the numbers under all bars and dividing by 12. Error bars show the standard error of the mean (for bars with more than 1 trial).*

on the 3 'hit' trials he directed many fixations to the background, whereas on the 'miss' trial he only fixated a single foreground object. This, together with the population data in Figure 4, appears to imply that, rather than looking at background colors, discounting them as potential targets and thus deciding not to report them (i.e. touching the sad face), the younger children often didn't look at the background when scanning the image and thus may literally not have seen background colors.

Overall, the results of Experiment 1 indicate that toddlers seek out color information in the immediately

perceivable environment mainly within discrete objects. Critically, the one-word color instruction, which should have focused participants' attention on color per se, proved to be insufficient to overcome this evidently very strong bias. If these data are indicative of a difficulty that toddlers experience in the conceptualization of the color domain specifically (and therefore in the learning of color words), it should be possible to eliminate the foreground–background distinction found in Experiment 1 by implementing the experimental task using a one-word *noun* instruction, whilst the region to be identified

Figure 4 *Eyetracking data on the Squares task of Experiment 1. The four panels show results for the different age groups, as labeled. Trials are separated by stimulus type (whether/where the target was present) and by response type (correct, incorrect). The three pairs of bars show the three stimulus types: trials on which the target was present in the foreground, present in the background, or absent, as labeled at the bottom of the figure. The two bars in each pair are for trials on which the subject responded correctly ('hit' or 'correct reject') or incorrectly; these are labeled COR and INC, respectively. The colors within each bar show the proportion of fixations directed to different regions of interest within the image. Gray: proportion of fixations to the background. Cyan: proportion of fixations to the foreground. Darker blue: on trials where the target was in the foreground, proportion of fixations to the target object. Fixations to the target are outlined in red. As described in the General Methods, fixations that were to the background but within 20 pixels of the edge of a foreground object were taken as being to the object. Fixations to the black region of the screen, or off-screen, were removed from the analysis. On each trial where more than 10 valid fixations were recorded, we calculated the proportion of fixations of each type. We then averaged this proportion across trials of each type, averaging across trials and subjects. The numbers along the bottom show the total number of trials contributing to each bar; these are slightly lower than in Fig 2 since on some trials fewer than 10 valid fixations were recorded. No one aged over 3 years made a mistake on a 'target present in foreground' trial, thus there are no data for this condition.*

within the visual stimulus is the same as that named by a color word. This possibility was investigated in Experiment 2.

Experiment 2

Experiment 2 further investigated whether toddlers' conceptual representation of color words is intrinsically associated with discrete entities. In this experiment, the stimuli were meaningful scenes depicting a house, a bus, a playground and an aquarium (see Table 2). In a

between-groups design, participants were required to locate a target named either by a color adjective (e.g. 'Blue!') or by a noun (e.g. 'Water!'). It was hypothesized that if knowledge of color words is limited, during the early years, to conceptual representations of the color domain that are inseparable from discrete entities, whereas knowledge of count and mass nouns is firmly established in terms of word–object mappings, or even more specifically in terms of auditory–visual mappings, then, children as young as 2 years of age should be able to locate a target region within a visual display without difficulty when cued with a noun, whether the region of interest maps onto the foreground or the background of a visual stimulus. Instead, when cued with a color word, background targets should be relatively hard to locate, despite mapping onto the same region named by a noun in the noun-cue condition (e.g. 'blue' vs. 'water').

Results and discussion

Accuracy and reaction time

Figure 6 summarizes accuracy scores across age groups on the Adjective (left) and Noun (right) conditions. As
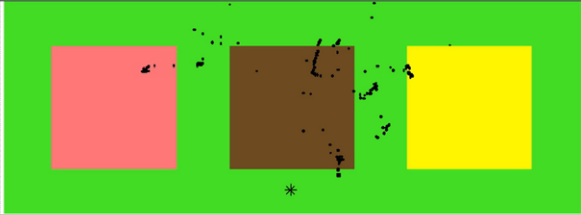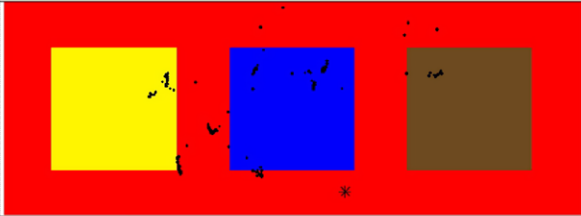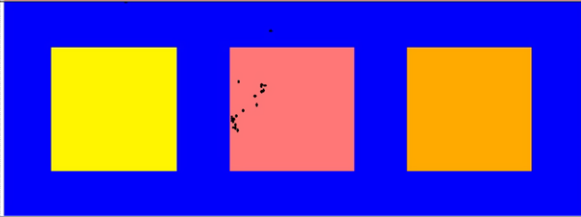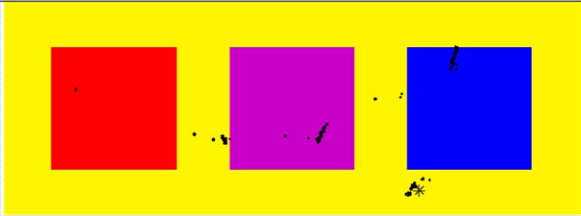
Color stimulus and fixation patterns

Trial	Target	Response	
2	green	hit	
4	red	hit	
7	blue	miss	
11	yellow	hit	

Figure 5 *'Background present' trials presented to subject 224. Unusually, this 2.6–year-old boy successfully identified a color presented in the background on 3 out of the 4 'target background' trials; the asterisk on the 'hit' trials indicates where on the image he touched. He failed to find the target 'blue' when it was present in the background (third row), although he successfully identified it when it was presented as a foreground square (not shown).*

done in Experiment 1, accuracy and RTs were averaged across the four targets for foreground, background and absent trials, respectively; data are given in Table 5. In the Adjective condition, despite using rather different visual stimuli, we largely replicated the results of Experiment 1 (Figure 2). Once again, a mixed-design Analysis of Variance revealed significant main effects of age group, $F(3, 45) = 15.003$, $p < .001$, and trial type (foreground/background/absent), $F(2, 90) = 21.009$, $p < .001$, and a significant interaction between age group and trial type, $F(6, 90) = 4.934$, $p < .001$. As in the Squares task, younger children were less likely to point to a target color if it was present in the background than if it was in the foreground. Within-group comparisons of accuracy on foreground vs. background trials ($p = [.05/$ 4] .013 for statistical significance) showed that the 2- to 3-year-olds performed significantly worse on background trials compared to foreground (31% vs. 85% correct), $t(11) = 4.290$, $p < .001$), but this difference was not significant in the other age groups (4-year-olds 65% vs. 95% correct, $t(9) = 2.250$, $p = .051$; 5- to 12-year-olds 92% vs. 98% correct, $t(11) = 1.000$ $p = .339$; adults 95% vs. 100% correct, $t(14) = 1.871$, $p = .082$).

Accuracy scores in the Noun condition were quantitatively and qualitatively different from those obtained in the Adjective condition. Analysis of Variance entering both age group and condition (Noun, Adjective) as the between-groups factors, showed the expected main effects of age group, $F(3, 90) = 26.553$, $p < .001$, and trial type, $F(2, 180) = 28.772$, $p < .001$, and the
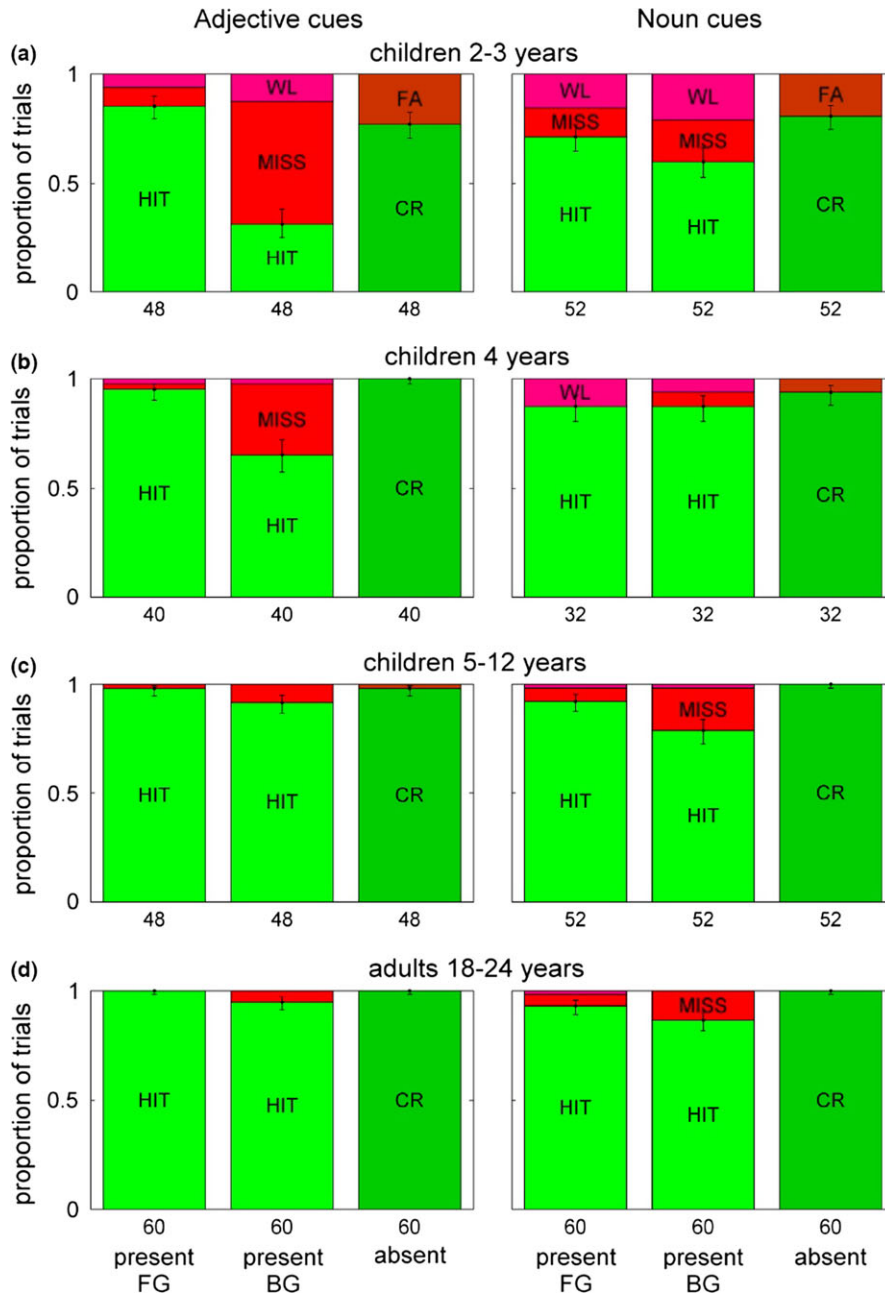
Figure 6   *Accuracy across age groups in Experiment 2: Scenes task. Left-hand panels: with Adjective cues ('red', 'yellow', 'blue', 'green'). Right-hand panels: with Noun cues (e.g. 'house', 'bus', 'window'; see Table 2). The three bars in each panel show, respectively, results for the four 'target present in the foreground' trials, the four 'target present in the background' trials, and the four 'target absent' trials. The four rows are for different age groups, as labeled. The adopted color-code is identical to that used in Figure 2.*

interaction between age group and trial type, $F(6, 180) = 3.253$, $p = .005$. The main effect of condition was not statistically significant, $F(1, 90) = 0.24$, $p = .878$, nor was the interaction between condition and age group, $F(3, 90) = 1.276$, $p = .288$. However, the interactions between

trial type and condition, $F(2, 180) = 4.326$, $p = .015$, and between age group, trial type and condition, $F(6, 180) = 3.280$, $p = .004$, were both statistically significant.

Irrespective of age group, accuracy on background trials was worse than on foreground trials for both the

Table 5 *Results of Experiment 2: Scenes. Reaction times are in seconds; standard error is given in parentheses*

Age group	Adjective cues				Noun cues			
	N subjects	Target present in fore-ground	Target present in back-ground	Target absent	N subjects	Target present in fore-ground	Target present in back-ground	Target absent
2–3-year-olds								
Number of trials:	12	48	48	48	13	52	52	52
Proportion of trials correct		0.85	0.31	0.77		0.71	0.60	0.81
Reaction time on correct trials		3.34 (0.19)	4.18 (0.37)	5.59 (0.21)		3.40 (0.17)	4.23 (0.26)	5.76 (0.42)
Reaction time on incorrect trials		3.77 (0.39)	5.16 (0.23)	4.63 (0.41)		6.26 (0.80)	5.54 (0.51)	4.82 (1.1)
4-year-olds								
Number of trials:	10	40	40	40	8	32	32	32
Proportion of trials correct		0.95	0.65	1.00		0.88	0.88	0.94
Reaction time on correct trials		2.27 (0.12)	2.74 (0.33)	3.36 (0.29)		3.45 (0.39)	2.99 (0.19)	4.33 (0.23)
Reaction time on incorrect trials		3.54 (0.29)	3.68 (0.28)	–		5.63 (1.26)	5.66 (2.13)	3.15 (0.97)
5–12-year-olds								
Number of trials:	12	48	48	48	13	52	52	52
Proportion of trials correct		0.98	0.92	0.98		0.92	0.79	1.00
Reaction time on correct trials		1.92 (0.07)	1.94 (0.07)	2.57 (0.09)		2.43 (0.20)	2.85 (0.28)	2.69 (0.15)
Reaction time on incorrect trials		5.26 (–)	2.58 (0.19)	2.84 (–)		4.67 (0.49)	3.92 (0.53)	–
Adults								
Number of trials:	15	60	60	60	15	60	60	60
Proportion of trials correct		1.00	0.95	1.00		0.93	0.87	1.00
Reaction time on correct trials		1.32 (0.03)	1.37 (0.05)	1.52 (0.03)		1.69 (0.05)	1.97 (0.16)	1.79 (0.05)
Reaction time on incorrect trials		–	2.46 (0.64)	–		3.64 (0.38)	2.07 (0.16)	–

Adjective condition (72% vs. 95%, $t(48) = 4.302$, $p < .001$) and the Noun condition (78% vs. 86%, $t(48) = 2.624$. $p = .012$, paired samples t-tests). The three-way interaction was followed up using eight further paired samples t-tests, ($p = [.05/8]$ .006 for statistical significance), comparing accuracy on foreground and background trials within age groups and separately for the two conditions. The latter showed that only the 2- to 3-year-olds performed significantly worse on background trials, and only in the Adjective condition (31% vs. 85%, $t(11) = 4.290$, $p = .001$).

Overall, when the target was present in the foreground, all age groups performed near perfectly, but on average slightly worse when the target was indicated with a noun than when it was indicated by its color (independent sample t-test across all ages: 86% (Noun) vs. 95% (Adjv), $t(96) = 2.882$, $p = .005$). This implies that participants found the task using noun cues harder than that using color cues. Reaction time (RT) data supported this conclusion. This is plotted in Figure 7. Analysis of Variance computed on RT data found significant main effects of condition, $F(1, 79) = 7.849$, $p < .01$, age group, $F(3, 79) = 59.668$, $p < .001$, and trial type, $F(2, 158) = 39.584$, $p < .001$, and a significant interaction between trial type and age group, $F(6, 158) = 10.565$, $p < .001$. The interactions between age group and condition, $F(3, 79) = 1.227$, $p > .05$, and between trial type and condition, $F(2, 158) = 1.372$, $p > .05$, were not statistically significant, whereas the interaction between age

group, trial type and condition approached significance, $F(6, 158) = 2.035$, $p = .064$. Overall, participants in the Noun condition took longer to respond on both foreground, $t(96) = 2.222$, $p = .029$, and background trials, $t(86) = 2.101$, $p = .039$, the difference being particularly strong and statistically significant among the adults (foreground: $p < .001$; background: $p = .001$), supporting the idea that the use of noun cues made the task slightly harder.

Arguably, our stimuli made the Noun condition particularly challenging in the background trials. In the 'yellow present in background' scene, for example (Table 2), participants were instructed to touch 'bus'. Their response was counted as correct only if they touched the yellow body of the bus; touching the window or door, for example, was counted as incorrect. This is one reason for the increased 'wrong location' errors visible in Figure 6. It is particularly striking, therefore, that when the target was present in the background, performance for the youngest age group was substantially better with noun cues than with adjective cues. Despite the fact that performance overall is worse with noun cues, in this group performance rose from just 31% correct with adjective cues to a respectable 60% correct with noun cues on background trials. With the adjective cue, performance in the youngest age group was significantly worse on background than on foreground trials (for 2–3-year-olds, $t(11) = 4.29$, $p = .001$; for 4-year-olds, $t(9) = 2.25$, $p = .051$), as found in Experiment 1. With the
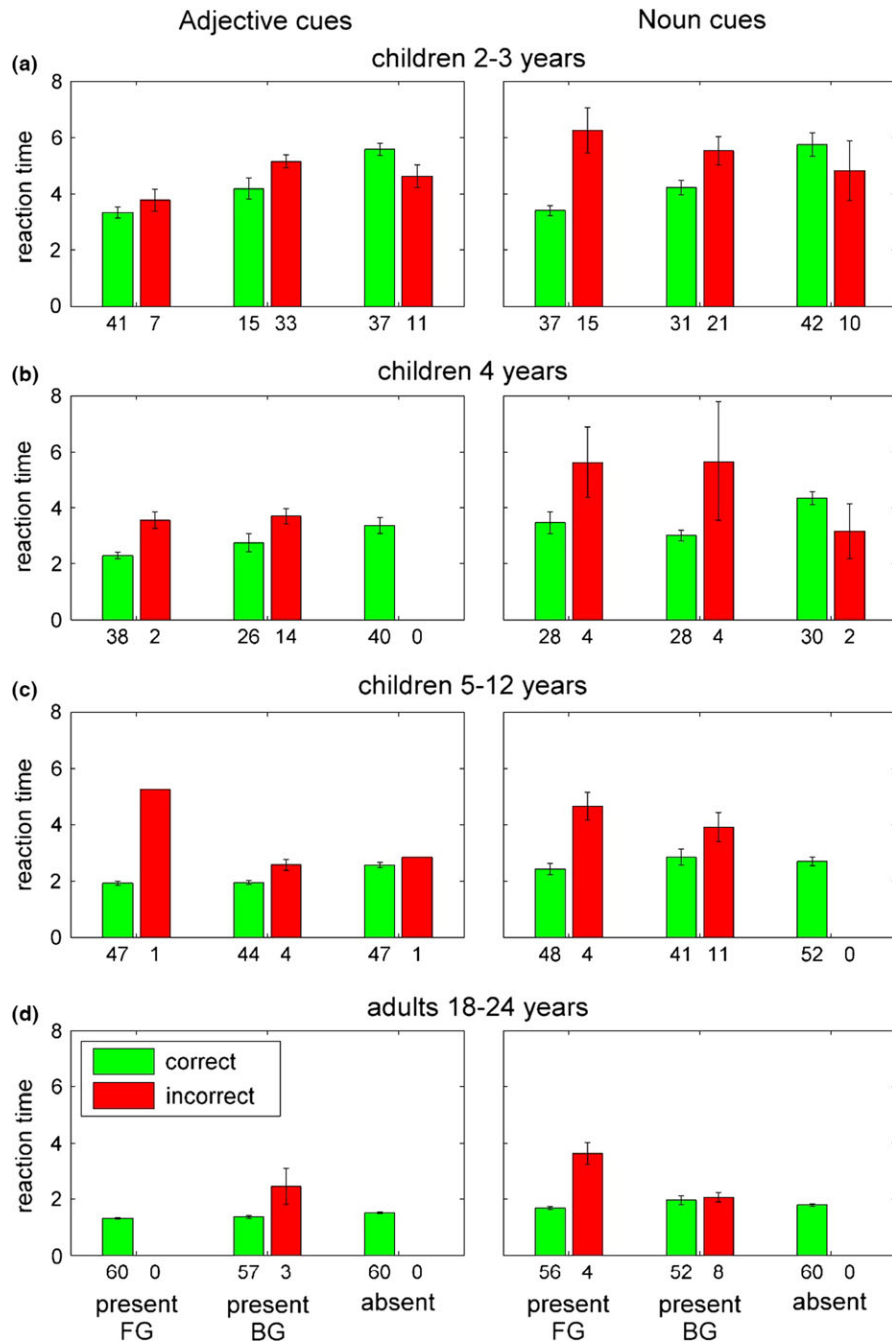
Figure 7   *Reaction times on the Scenes task of Experiment 2. Left-hand panels: with Adjective cues ('red', 'yellow', 'blue', 'green'). Right-hand panels: with Noun cues (e.g. 'house', 'bus', 'window'; see Table 2). Other details as per Figure 3.*

noun cue, no such difference existed for any age group (for 2–3-year-olds, $t(12) = 1.389$, $p = .190$; for 4-year-olds, $t(7) = 0.00$, $p = 1$).

Examining the Scenes stimuli, shown in Table 2, there is a clear difference between scenes where 'red' or 'yellow' was the target color in the background, and 'blue'/'green' target backgrounds. In the latter, the

background color fills the figure, and is identifiable by the knowledge that grass is usually green and water conventionally blue. In the former, the background does not completely fill the figure, but has a shape which enables it to be identified despite the lack of any necessary relation between the target color and the object depicted (houses are not always red, buses are not

always yellow). One might therefore expect a difference in performance between the 'house/bus' and 'aquarium/playground' Noun trials. As Figure 8 shows, the youngest age-group was indeed less successful in finding a Noun target when it was present in the house/bus scenes, as compared to the aquarium/playground scenes (p=0.04, paired t-test comparing accuracy on all 8 "target present" trials for 2–3 year olds in house/bus scenes vs aquarium/playground scenes; the difference is not significant when comparing only the 4 foreground or 4 background trials). But critically, the effect of scene type applied whether the target was in the foreground or background. Paired samples t tests comparing accuracy on foreground vs background trials (2–3 year olds) within a scene-type showed no significant differences (both $p > .05$). In other words, shifting to Noun cues removes the difference between background and foreground trials which we found with Adjective cues in this age-group and this occurs in both classes of scene that we used.

Eye tracking

Figure 9 shows the eye tracking data for the Scenes task. The pattern of results is very similar to that shown in Figure 4 for the Squares task. Once again, relatively more fixations were directed towards the target on trials on which the target was successfully identified. As before, Analysis of Variance found a significant effect of age group, $F(3, 379) = 11.953$, $p < .001$, and trial type, $F(2, 379) = 58.541$, $p < .001$; while the interaction

between age group and trial type was not significant, $F(6, 379) = 0.779$, $p = .586$ (three-way ANOVA on the proportion of fixations to background on correctly answered trials only, with the factors being age group, trial type (foreground, background or absent) and cue (Adjective or Noun)). Critically, there is no significant difference between the Adjective and Noun conditions (main effect of cue, $F(1, 379) = 0.002$, $p = .963$). This indicates that the very different performance observed with adjective vs. noun cues was not mediated by differences in the patterns of fixations. Children tended to spend the same amount of time fixating the background in both cases, but they were much less likely to detect a background target if they were cued with a color adjective rather than with a noun.

Overall, the Adjectives condition of Experiment 2 replicates Experiment 1 with a different stimulus, confirming that participants find it harder to locate target colors presented as ground. The Nouns condition demonstrates that this impairment depends on the nature of the linguistic cue. Merely changing the cue from a color word to a concrete noun nearly equalizes performance for foreground vs. background targets, removing the bias found in Experiment 1.

General discussion

The two experiments reported here used a visual search task with children aged 2 to 12 years, and with adults. Participants were shown colorful visual displays on a
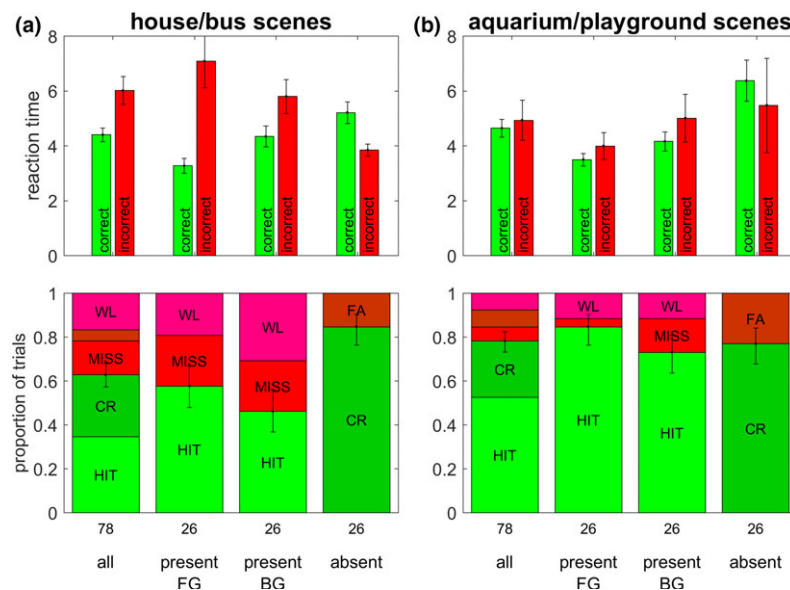

Figure 8 *Comparing performance on house/bus vs. aquarium/playground scenes in the Noun conditions, youngest age group only (2–3-year-olds). AB: Reaction times; CD: proportion correct. Details as in Figures 2 and 3.*
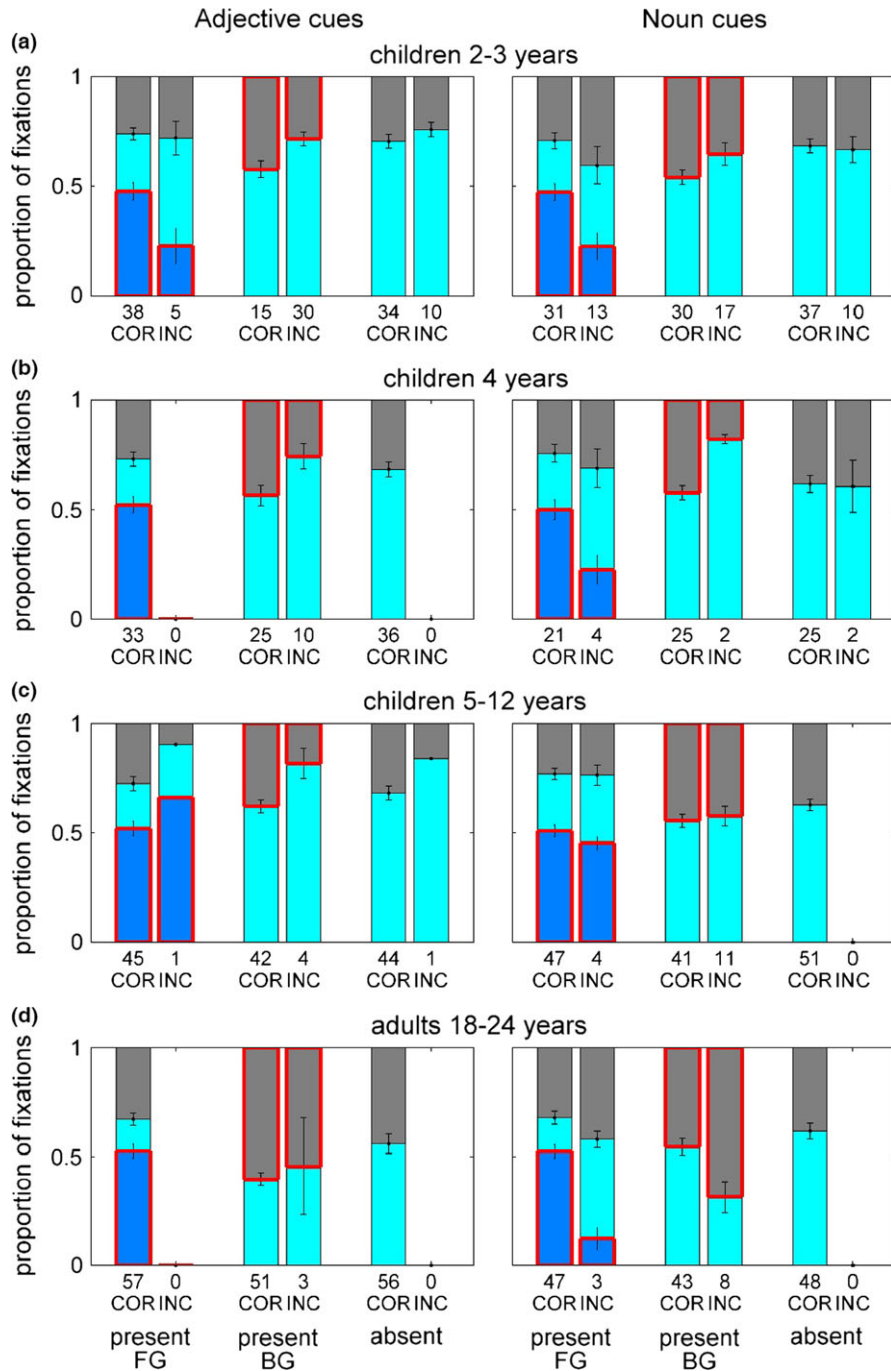
Figure 9 *Eye tracking data for Experiment 2: Scenes task. Left: Adjectives condition; Right: Nouns condition. The different rows are for different age groups. Other details as per Figure 4.*

touchscreen monitor, and were required to point to a target location after being cued either with an adjective (e.g. 'Blue!') or with a noun (e.g. 'Water!'). Targets were present either in the foreground or in the background of the visual displays. Foreground targets were either

abstract objects such as a square (Experiment 1), or meaningful objects such as a door (Experiment 2). Background targets were either a uniform background (Experiment 1), or depicted meaningful and known entities such as water or a bus (Experiment 2). Data were

analyzed in four developmental age groups: 2- to 3-year-olds, 4-year-olds, 5- to 12-year-olds, and adults. In terms of development, the two younger age groups are typically assessed separately in developmental research, in line with the notable cognitive and linguistic development that takes place at this age.

We report that when toddlers are asked to locate a known color within either an abstract or a meaningful visual scene, they will often fail to do so unless the target color is presented within the context of a discrete object, whether an abstract shape or a familiar object. When a target color is presented as the background of the image, it is highly likely that 2- to 3-year-olds, and to a lesser extent 4-year-olds, will not find it, even though they are near-perfect when the same color is presented in the foreground. The deficit remains even if the background color is part of a meaningful scene, e.g. blue water in an aquarium, if children are cued to look simply for the color ('blue') rather than for an object or substance ('water'). This 'blindness to background' effect persisted to some extent in older children and even in adults. In toddlers, it is a striking phenomenon which has not previously been reported. We suggest that it reflects an inbuilt tendency of humans to attend to objects, which may assist young children in acquiring language.

Color is an important aspect of human visual experience. Yet young toddlers take many months of training and reinforcement before they learn to apply color words correctly and consistently (Rice, 1980; Roberson *et al.*, 2004). This indicates a fundamental difficulty during the early years in the development of color concepts, and raises important questions when one considers that children's environments are rich in color both perceptually and linguistically.

Previous studies that have examined understanding of the color domain during the early years have done so by implementing non-verbal tasks, such as object matching, or verbal tasks, such as teaching color words. By and large, these studies have shown that the development of color concepts and the acquisition of color words unfold in atypical fashion relative to the manner in which concepts are formed and words are learned in other, lexically similar categories such as shape and size. Importantly, the majority of studies that have investigated young children's understanding of the color domain, both conceptually (Kowalski & Zimiles, 2006; Sandhofer & Smith, 1999; Soja, 1994) and linguistically (Baldwin, 1989; Gottfried & Tonks, 1996; O'Hanlon & Roberson, 2006, 2007), reported a tendency to direct attention to the shape of objects. For example, Soja (1994) showed that 2-year-olds are able to match objects by color only when these are identical (e.g. cups) or highly similar (e.g. four-legged animals) in overall shape.

On the other hand, word extension tasks in which shape, texture or color were viable word interpretations showed a tendency to use shape information to establish word meaning beyond a named object (e.g. Baldwin, 1989). Children also prioritize matching on the basis of semantic content rather than color (Mani, Johnson, McQueen & Huettig, 2013); for example, if primed with the word 'banana', toddlers would look more rapidly to another edible object (e.g. a cookie) than to another yellow object (e.g. a yellow cup). This tendency to selectively attend to object shape may hinder the ability to focus attention on color. In consequence, children would find color words hard to learn, even when they can implicitly use information about object color (Johnson & Huettig, 2011; Johnson *et al.*, 2011).

This study took a novel approach to investigate whether toddlers struggle to attend to color in isolation from objects. Four aspects of our procedure were critical to address this question. First, by asking participants to locate a *known* color during a visual search task, we were able to establish whether directing attention to color is problematic even when conceptual and lexical color knowledge are in place from the outset. Second, we trained participants to respond to a one-word instruction, a color name, which maximized the likelihood of focusing attention on color per se, and eliminated the possibility of a syntax bias influencing the results. Third, we manipulated whether color targets were figure or ground, and made the ground area within the visual stimulus much larger than that occupied by the foreground objects. This should have maximized the likelihood of seeing background colors even in the younger children, especially considering that the stimuli were only 50 cm away from their forehead. Fourth, we recorded eye movements during visual search, which enabled us to examine the proportion of fixations to the foreground and to the background of visual displays in children as young as 2 years of age, relative to the location of named targets. In addition, we tested children at three stages of development – young toddlers (2–3-year-olds), those at the end of toddlerhood (4-year-olds), and children up to the beginning of adolescence (5–12-year-olds). To examine lifespan developmental trends, adults were also tested on the same tasks implemented with children. Previous studies have demonstrated a 'shape bias' in 3-year-olds as well as in adults (Landau, Smith & Jones, 1992b).

Huettig *et al.* (Huettig, Olivers & Hartsuiker, 2011; Huettig, Mishra & Olivers, 2012) have argued that working memory is critical in language-mediated visual search, binding together the information about word meaning stored in long-term memory with visual attentional orienting responses. One might expect young children to have a stronger mapping from concrete nouns

(e.g. 'house') to their referents than from adjectives (e.g. 'red'). In fact, reaction times were overall slightly longer in the Noun condition than in the Adjective condition (Figure 7), which does not suggest a stronger mapping. The most striking effect was the inaccurate performance when target colors were cued with an adjective and presented in the background of the image. This suggests that although children understood the concept of 'red' well, their visual search mechanisms tended to default to searching specifically for red *objects*, ignoring red when this formed the background of a scene.

One reason for this blindness to background must be that children simply spent less time fixating the background. In the Squares task, target-absent trials, children directed only around 30% of fixations towards the background, although this made up 63% of the stimulus area. On target-background trials, successful detection was associated with more fixations directed towards the background. However, this cannot be the whole explanation. As we saw in the eye tracking data displayed in Figure 9, children spent similar proportions of time viewing the background in the Adjective and the Noun conditions, but were much more likely to identify a background target in the latter. Thus, toddlers find it easier to identify targets in the background of visual scenes when these are cued with a noun; for example, they are looking for 'grass' rather than 'green' in our scenes; see Table 2. We interpret this as a form of object bias: children find it easier to direct attention to objects identified by color – even when the object in question is uncountable material like grass rather than a count noun like a tree – than to color as an abstract concept. Johnson *et al.* (2011) showed that children can use object color information, such as their stored knowledge that frogs are green, before they are able to demonstrate knowledge of the word 'green'. Our work suggests that even when children have learnt basic color terms and can successfully identify 'the green object', they still struggle with the concept of *green* in isolation from objects.

In addition, and importantly, the finding that toddlers (and adults) tended to search for color targets within the context of figure not ground, regardless of whether the figure depicted a shape or a meaningful object and, critically, whilst searching for targets they had good knowledge of, would suggest that the 'shape bias' documented in word learning studies might be an extension of a more fundamental cognitive mechanism to attend to discrete entities in one's environment, rather than a 'strategy' specific to word learning that develops over the first 2 years of life specifically to guide lexical acquisition (e.g. Colunga & Smith, 2005; Smith, 2000). Our data suggest that the whole object bias (Woodward & Markman, 1998) and the shape bias (Landau *et al.*,

1988) might reflect the same process, and that the findings of the present study were driven by perceptual and attentional predispositions that may well be independent of word learning strategies. In line with these ideas, recent studies have shown that the shape bias in lexical acquisition studies might be task induced (e.g. Cimpian & Markman, 2005; Diesendruck & Bloom, 2003).

Taken as a whole, our results suggest that toddlers and even adults tend to direct their attention towards things they perceive as particular entities. Usually, this means attending to figure rather than ground in the classic Gestalt sense. However, our data show that even the youngest children can successfully attend to ground when given the appropriate cue. When lexical meaning was provided for the ground (e.g. 'water' or 'grass'), toddlers directed attention to it as successfully as did older children (and adults). The ability to switch attention from figure to ground is impaired specifically when *color* is the (visual and lexical) target, and to a much greater extent in 2- to 4-year-olds compared to older children. This suggests that color conceptualization is still largely developing during toddlerhood whilst it is firmly established in later childhood. Our data are consistent with the possibility that, at a young age, children don't conceptualize color as existing independently of objects (while even for adults, there is a tendency for color terms to direct attention towards objects). Thus, inattention to color during the early years may be the primary reason for children's slow and effortful development of color concepts and acquisition of color words.

An important caveat is that most of the studies we have cited above were carried out in Western countries. In our study, all participants were white, middle-class, native-English-speaking Caucasians living in the North East of England. We do not know whether the 'blindness to background' effect would have been seen in other groups. There is wide-ranging evidence suggesting that, compared to Westerners, East Asian participants tend to report more holistic perceptions (Abel & Hsu, 1949) and are more affected by the background or setting of a target (Park, Nisbett & Hedden, 1999; Ji, Peng & Nisbett, 2000). Masuda and Nisbett (2001) report that, shown a picture reminiscent of our own 'aquarium' scene, American participants tended to begin by reporting the foreground objects ('there was what looked like a trout swimming to the right'), whereas Japanese participants generally began by describing the background ('I saw what looked like a stream; the water was green.'). Western toddlers also hear more noun phrases while East Asian toddlers hear more verbs (Choi & Gopnik, 1995; Tardif, Shatz & Naigles, 1997). In general, Nisbett, Peng,

Choi and Norenzayan (2001) state that, 'East Asians would be expected to attend more to the field than European Americans, who would be expected to attend more to a salient target object.' If so, one might speculate that East Asian children would show less 'blindness to background' than the English children reported here.

In sum, the current study contributes two novel findings to the developmental literature. First, when young, English-speaking toddlers are cued with a known color adjective and are required to locate the corresponding color target present in the background of a simple visual display, they by and large fail to do so, showing a 'blindness to background' effect. Eye tracking data showed that this effect is partially mediated by a tendency to direct attention to discrete objects. However, when cued with a count or mass noun, children at all ages (and adults) easily switch attention from figure to ground (e.g. when locating 'water' rather than 'blue'). Second, eye movement patterns did not differ between adjective and noun conditions, demonstrating that the 'blindness to background' effect is mediated by the nature of the linguistic cue, rather than an inability to switch attention from figure to ground. Taken together, these findings indicate a fundamental incompleteness of the semantic representation of color concepts at a young age, with these being inseparable from objects.

In addition to these two critical findings, the results of our visual search task suggest that the whole object and shape biases reported in the developmental literature may result from a common cognitive function, grounded in attention, that predisposes humans to attend to figure not ground when interpreting language. In line with a large body of previous developmental research, we agree that this cognitive predisposition supports and facilitates the complex and crucially important task of lexical acquisition during the early years.

Acknowledgements

References

Abel, T.M., & Hsu, F.L. (1949). Some aspects of personality of Chinese as revealed by the Rorschach Test. *Rorschach Research Exchange and Journal of Projective Techniques*, **13** (3), 285–301.

Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science*, **8**, 199–209. doi:10.1111/j.1467-7687.2005.00406

Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development*, **67**, 635–645. doi:10.1111/j.1467-8624.1996.tb01756

Au, T.K., & Laframboise, D.E. (1990). Acquiring colour names via linguistic contrast: the influence of contrasting terms. *Child Development*, **61**, 1808–1823.

Au, T.K., & Markman, E.M. (1987). Acquiring word meaning via linguistic contrast. *Cognitive Development*, **2** (3), 217–236. doi:10.1016/S0885-2014(87)90059-1

Baldwin, D.A. (1989). Priorities in children's expectations about object label reference: form over colour. *Child Development*, **60** (6), 1291–1306.

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition*, **127** (6), 391–397. doi:0.1371/ journal.pone.0073359

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley, CA: University of California Press.

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

Carey, S. (1978). The child as a word learner. In M. Halle, J. Bresnan & G.A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.

Choi, S., & Gopnik, A. (1995). Early Acquisition of Verbs in Korean: a cross-linguistic study. *Journal of Child Language*, **22** (3), 497–529.

Cimpian, A., & Markman, E.M. (2005). The absence of a shape bias in children's word learning. *Developmental Science*, **41** (6), 1003–1019. doi:10.1037/0012-1649.42.6.1339

Colunga, E., & Smith, L.B. (2005). From the lexicon to expectations about kinds: a role for associative learning. *Psychological Review*, **112** (2), 347–382. doi:10.1037/0033-295X.112.2.000

Diesendruck, G., & Bloom, P. (2003). How specific is the shape bias? *Child Development*, **74** (1), 168–178. doi:10.1111/1467-8624.00528

Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., & Thal, D., et al. (1994). Variability in early communicative development.

Monographs of the Society for Research in Child Development, **59** (5, Serial No. 242).

Gershkoff-Stowe, L., & Smith, L.B. (2004). Shape and the first hundred nouns. *Child Development*, **75** (4), 1098–1114. doi:10.1111/j.1467-8624.2004.00728

Gliga, T., Mareschal, D., & Johnson, M.H. (2008). Ten-month-olds' selective use of visual dimensions in category learning. *Infant Behavior and Development*, **31**, 287–293.

Golinkoff, R.M., Hirsh-Pasek, K., Bloom, L., Woodward, A.L., Akhtar, N., et al. (2000). *Becoming a word learner: A debate on lexical acquisition*. New York: Oxford University Press.

Gottfried, G.M., & Tonks, J.M. (1996). Specifying the relation between novel and known: input affects the acquisition of novel color terms. *Child Development*, **67**, 850–866.

Henderson, J.M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, **7** (11), 498–504.

Huettig, F., Mishra, R.K., & Olivers, C.N.L. (2012). Mechanisms and representations of language-mediated visual attention. *Frontiers in Psychology*, **2**, 394.

Huettig, F., Olivers, C.N.L., & Hartsuiker, R.J. (2011). Looking, language, and memory: bridging research from the visual world and visual search paradigms. *Acta Psychologica*, **137** (2), 138–150.

Ji, L.J., Peng, K., & Nisbett, R.E. (2000). Culture, control, and perception of relationships in the environment. *Journal of Personality and Social Psychology*, **78** (5), 943–955.

Johnson, E.K., & Huettig, F. (2011). Eye movements during language-mediated visual search reveal a strong link between overt visual attention and lexical processing in 36-month-olds. *Psychological Research*, **75** (1), 35–42.

Johnson, E.K., McQueen, J.M., & Huettig, F. (2011). Toddlers' language-mediated visual search: they need not have the words for it. *Quarterly Journal of Experimental Psychology*, **64** (9), 1672–1682.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., & Murray, R. et al. (2007). What's new in Psychtoolbox-3? *Perception*, **36**, 14 (European Conference on Visual Perception Abstract Supplement).

Kowalski, K., & Zimiles, H. (2006). The relation between children's conceptual functioning with colour and colour term acquisition. *Journal of Experimental Child Psychology*, **94**, 301–321.

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

Landau, B., Smith, L.B., & Jones, S.S. (1992a). Object shape, object function, and object name. *Journal of Memory and Language*, **38** (1), 1–27.

Landau, B., Smith, L.B., & Jones, S.S. (1992b). Syntactic context and the shape bias in children's and adults' lexical learning. *Journal of Memory and Language*, **31** (6), 807–825.

McDonough, L. (2002). Early concepts and early language acquisition: what does similarity have to do with either? In N.L. Stein, P. Bauer & M. Rabinowitz (Eds.), *Representation,* *memory, and development: Essays in honor of Jean Mandler* (pp. 115–143). Mahwah, NJ: Erlbaum.

Mandler, J.M. (2004). Thought before language. *Trends in Cognitive Sciences*, **8** (11), 508–513.

Mani, N., Johnson, E.K., McQueen, J.M., & Huettig, F. (2012). How yellow is your banana? Toddlers' language-mediated visual search in referent-present tasks. *Developmental Psychology*, **49** (6), 1036–1044. doi:10.1037/a0029382

Markman, E.M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Masuda, T., & Nisbett, R.E. (2001). Attending holistically versus analytically: comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology*, **81** (5), 922–934.

Memmert, D. (2006). The effects of eye movements, age, and expertise on inattentional blindness. *Consciousness and Cognition*, **15** (3), 620–627. doi:10.1016/j.concog.2006.01.001

Nisbett, R.E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: holistic versus analytic cognition. *Psychological Review*, **108** (2), 291–310.

O'Hanlon, C.G., & Roberson, D. (2006). Learning in context: linguistic and atttentional constraints on children's colour term learning. *Journal of Experimental Child Psychology*, **94**, 275–300.

O'Hanlon, C.G., & Roberson, D. (2007). What constrains children's learning of novel shape terms? *Journal of Experimental Child Psychology*, **97**, 138–148.

Park, D.C., Nisbett, R., & Hedden, T. (1999). Aging, culture, and cognition. *Journal of Gerontology B*, **54**, 75–84.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, **10** (4), 437–442.

Pitchford, N.J., & Mullen, K.T. (2001). Conceptualization of perceptual attributes: a special case for color? *Journal of Experimental Child Psychology*, **80**, 289–314.

Remington, A., Cartwright-Finch, U., & Lavie, N. (2014). I can see clearly now: the effects of age and perceptual load on inattentional blindness. *Frontiers in Human Neuroscience*, **8**, 229. doi:10.3389/fnhum.2014.00229

Rice, M. (1980). *Cognition to language categories: Word, meaning and training*. Baltimore, MD: University Park Press.

Roberson, D., Davidoff, J., Davies, I.R.L., & Shaprio, L.R. (2004). The development of colour categories in two languages: a longitudinal study. *Journal of Experimental Psychology: General*, **133**, 554–571.

Sandhofer, C.M., & Smith, L.B. (1999). Learning colour words involves learning a system of mappings. *Developmental Psychology*, **35**, 668–679.

Simons, D.J. (2000). Attentional capture and inattentional blindness. *Trends in Cognitive Sciences*, **4** (4), 147–155.

Smith, L.B. (2000). Learning how to learn words: an associative crane. In R.M. Golinkoff, K. Hirsh-Pasek, L. Bloom, L.B. Smith, A.L. Woodward, et al. (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 61–80). Oxford: Oxford University Press.

Soja, N.N. (1994). Young children's concept of colour and its relation to the acquisition of colour words. *Child Development*, **65** (3), 918–937. doi:10.1111/j.1467-8624.1994.tb00793

Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: a comparison of English, Italian, and Mandarin. *Journal of Child Language*, **24** (3), 535–565.

Waxman, S.R. (2002). Links between object categorization and naming: origins and emergence in human infants. In D.H. Rakison & L.M. Oakes (Eds.), *Early category and concept development: Making sense of the blooming, buzzing confusion* (pp. 213–241). New York: Oxford University Press.

Waxman, S.R., & Markow, D.B. (1998). Object properties and object kind: twenty-one-month-old infants' extension of novel adjectives. *Child Development*, **69** (5), 1313–1329. doi:10.1111/j.1467-8624.1998.tb06214

Woodward, A.L., & Markman, E.M. (1998). Early word learning. In W. Damon, D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology, Volume 2: Cognition, perception and language* (pp. 371–420). New York: John Wiley and Sons.