

*Manuscript in press at Cortex*

**The link between reading ability and visual spatial attention across development**

Alex L. White<sup>1,2\*</sup>, Geoffrey M. Boynton<sup>3</sup> & Jason D. Yeatman<sup>1,2</sup>

<sup>1</sup> Institute for Learning & Brain Sciences, University of Washington, Seattle

<sup>2</sup> Department of Speech & Hearing Sciences, University of Washington, Seattle

<sup>3</sup> Department of Psychology, University of Washington, Seattle

\*Corresponding author:

Alex White  
Department of Speech & Hearing Sciences  
Institute for Learning & Brain Sciences  
University of Washington  
1715 Columbia Rd NE, Box 357988  
Seattle, WA 98195  
(206) 685-9365  
email: alexlw@uw.edu.

*Note: the **Supplemental Material** referred to in the manuscript has been appended to the end of this document.*

## Abstract

Interacting with a cluttered and dynamic environment requires making decisions about visual information at relevant locations while ignoring irrelevant locations. Typical adults can do this with covert spatial attention: prioritizing particular visual field locations even without moving the eyes. Deficits of covert spatial attention have been implicated in developmental dyslexia, a specific reading disability. Previous studies of children with dyslexia, however, have been complicated by group differences in overall task ability that are difficult to distinguish from selective spatial attention. Here, we used a single-fixation visual search task to estimate orientation discrimination thresholds with and without an informative spatial cue in a large sample (N=123) of people ranging in age from 5 to 70 years and with a wide range of reading abilities. We assessed the efficiency of attentional selection via the *cueing effect*: the difference in log thresholds with and without the spatial cue. Across our whole sample, both absolute thresholds and the cueing effect gradually improved throughout childhood and adolescence. Compared to typical readers, individuals with dyslexia had higher thresholds (worse orientation discrimination) as well as smaller cueing effects (weaker attentional selection). Those differences in dyslexia were especially pronounced prior to age 20, when basic visual function is still maturing. Thus, in line with previous theories, literacy skills are associated with the development of selective spatial attention.

*Key words: spatial attention; visual perception; dyslexia; development*

## 1. Introduction

### 1.1 Spatial attention and reading ability

Complex and cluttered environments pose a challenge when using vision to navigate and interact: only a subset of the information entering the eyes is relevant to the task at hand. Fortunately, the brain is equipped with multiple mechanisms – generally referred to as “attention” – that prioritize relevant information and filter out irrelevant information. Selective spatial attention prioritizes information at specific locations in the scene, and is critical for everyday tasks. A prominent example is reading: pages of text are extremely cluttered, and letters are only identifiable in the central visual field (Legge et al., 2007). Even within the central visual field, it is difficult to fully process more than one word at a time (White, Palmer, & Boynton, 2018). Therefore, reading requires rapid shifts of spatial attention to select individual words on a page (Rayner, 2009).

Given its importance in reading, one might hypothesize that a deficiency in spatial attention would cause difficulty in reading. In fact, many researchers have argued for an association between spatial attention and developmental dyslexia, a reading disability that affects 5-10% of the population (e.g., Bosse, Tainturier, & Valdois, 2007; Facoetti, Paganoni, Turatto, Marzola, & Mascetti, 2000; Franceschini, Gori, Ruffino, Pedrolli, & Facoetti, 2012; Vidyasagar & Pammer, 2010). Some authors link the attentional deficits to a more general abnormality in the “dorsal visual stream” (e.g., Pammer, Hansen, Holliday, & Cornelissen, 2006; Vidyasagar & Pammer, 1999), based on performance in tasks thought to rely on the magnocellular visual pathway (Boden & Giaschi, 2007; Demb, Boynton, Best, & Heeger, 1998; Demb, Boynton, & Heeger, 1997; Simone Gori, Seitz, Ronconi, Franceschini, & Facoetti, 2015; Kevan & Pammer, 2008). However, many other etiological explanations for dyslexia have been proposed, involving the auditory system, phonological processing, or higher-level language skills (reviewed by Vellutino, Fletcher, Snowling, & Scanlon, 2004).

The goal of the present investigation was to evaluate the link between spatial attention and reading ability over development. Methods for measuring spatial attention have been

inconsistent across studies of children and adults with dyslexia. That is important because visual spatial attention comes in several forms: (1) *Overt* spatial attention involves movements of the head and eyes to align the high-resolution fovea with task-relevant objects. While reading this page, for example, you are making roughly four rapid eye movements per second (Rayner, 2009). (2) *Covert* spatial attention is the selective prioritization of locations in the visual field without moving the eyes. There are two types of covert spatial attention: (a) *Endogenous* attention is voluntary and driven by knowledge and goals. While fixating on the word “endogenous” in the previous sentence, your attention shifted to the next word, “attention”, before your eyes moved on (Rayner, 2009). (b) *Exogenous* attention is involuntary, triggered by salient external events that may or may not be task-relevant, like an email notification appearing in the corner of the screen while you try to read a manuscript.

Although there is some evidence for abnormal eye movement control (overt attention) in dyslexia (e.g., Eden, Stein, & Wood, 1994; Hawelka & Wimmer, 2005; Rayner, 1985), **here we focus on covert spatial attention. During reading, endogenous covert attention** is required to begin processing parafoveal words before the eyes move, to plan the saccades themselves, and perhaps even to shift attention between letters within fixated words (Rayner, 2009; Vidyasagar & Pammer, 2010).

In the laboratory, covert spatial attention is often studied by requiring an observer to fixate and respond to target stimuli presented in the peripheral visual field. Prior to the stimuli, a *cue* directs attention to one or more locations (Posner, 1980). **A cue could be an arrow that points to the location most likely to contain task-relevant stimuli, or a small shape that draws attention by flashing near a potential stimulus location.** The effects of covert spatial attention include increased discrimination accuracy and decreased response time when the target stimulus’s location is cued compared to uncued (reviewed by Carrasco, 2011). Physiologically, neural responses are greater for stimuli at cued locations than at uncued locations (Maunsell, 2015; Reynolds & Chelazzi, 2004).

In this study, we are interested in how covert spatial attention differs between individuals with and without dyslexia, in childhood as well as in adulthood. Because our goal is to measure visual task performance across a wide age range, we must first consider more general developmental changes in visual perception and attention.

## 1.2 The development of covert spatial attention

Since the 1980s, psychologists have studied the development of covert spatial attention using spatial cueing paradigms (Akhtar & Enns, 1989; Brodeur & Boden, 2000; Brodeur & Enns, 1997; Enns & Brodeur, 1989; Iarocci, Enns, Randolph, & Burack, 2009; Leclercq & Siéroff, 2013; Pearson & Lane, 1990; Plude, Enns, & Brodeur, 1994; Ristic & Enns, 2015; Ristic & Kingstone, 2009; Wainwright & Bryson, 2002). Nearly all such studies assess spatial attention by comparing reaction times (RTs) across different cue conditions (e.g., valid vs. invalid). The question is how the differences in RT, which index attention effects, change across development. There is a general consensus that exogenous (automatic, stimulus-driven) cueing effects are present from at least pre-school age and are stable through the lifespan. Endogenous (voluntary, top-down) cueing effects show more gradual developmental change, suggesting an increase in strategic control over spatial attention.

Beyond that, there is little agreement on the details of the time course and which internal mechanisms are changing. Some studies claim that endogenous attention becomes “adult-like” by age 10 (Goldberg, Maurer, & Lewis, 2001; Landry, Johnson, Fleming, Crewther, & Chouinard, 2019; Michael, Lété, & Ducrot, 2013; Ristic & Enns, 2015; Wainwright & Bryson, 2005), but others suggest later maturation (Brodeur & Enns, 1997; Leclercq & Siéroff, 2013; Schul, Townsend, & Stiles, 2003).

A general challenge in this type of study is to separate developmental change in a specific mechanism of attention from developmental change in overall task performance. Younger participants tend to respond slower and less accurately to the same stimulus as older participants. Does a valid-invalid cueing effect of 50 ms reflect the same degree of attentional modulation in a 5-year-old as it does in a 20-year old? It is difficult to say, especially if their

accuracies differ. Several previous studies have attempted to address this difficulty, for instance by normalizing RTs across age groups (Gaspelin, Margett-Jordan, & Ruthruff, 2015; Goldberg et al., 2001). It remains an open question, however, whether the mechanisms of selective spatial attention develop independently of basic visual sensitivity and task performance.

Moreover, cueing effects on reaction times could have many underlying causes: change in the quality of the sensory representation, the speed of evidence accumulation, response bias, and/or motor preparation. Only a handful of developmental studies have focused on how spatial cues affect detection or discrimination accuracy (Akhtar & Enns, 1989; Brodeur & Boden, 2000; Schul et al., 2003), which can give more information about the underlying mechanisms. Even then, it is difficult to compare cueing effects in units of proportion correct across groups that differ in their absolute levels of accuracy. The conclusion depends on an interaction between cue condition and age. Such interactions are difficult to interpret if the measurement (e.g., differences in proportion correct) may not be linearly related to the theoretical variable of interest (e.g., the effect of attention on visual processing; Loftus, 1978).

### 1.3 A paradigm to study visual attention in dyslexia and across development

Studies that compare spatial attention across good and poor readers face the same challenges as studies that compare across age groups. Many in the past have not precisely specified what differs in dyslexic individuals: overall visual sensitivity, motor ability, overt eye movements, or endogenous covert selection, etc. Many rely on RTs, and the participant groups often differ in their overall performance level (e.g., people with dyslexia can have slow processing speed in general; Peterson & Pennington, 2015) as well as any putative attention effects.

One research team overcame those challenges by measuring orientation discrimination thresholds in a simple visual search task with and without an informative spatial pre-cue (Roach & Hogben, 2004, 2007, 2008). The participants' task was to report the tilt direction (clockwise or counterclockwise from vertical) of a single Gabor stimulus that was presented along with a variable number of vertical distractors, all equidistant from fixation (Baldassi & Burr, 2000). The display was presented briefly enough to avoid eye movements to the stimuli,

and the target's location varied randomly from trial to trial. On some trials, a 100% valid spatial cue (a small dot) flashed near the target's location, just prior to the stimulus array. In each condition, the experimenters used a staircase procedure to estimate each participant's orientation discrimination threshold: the degrees of tilt required to perform the task with ~75% correct. Without the spatial cue, thresholds rise with increasing set size, because the internal representation of each stimulus is noisy and each distractor could be mistaken for the target. With the spatial cue, thresholds in control participants are much lower (better performance) and less affected by increasing set size. This difference in threshold represents the benefit of spatial attention: the cue allows the participant to base their decision primarily on information at the target's location and exclude noise from the distractors (Roach & Hogben, 2007, 2008).

Roach & Hogben found that although adults with dyslexia performed normally in the uncued condition, their thresholds were abnormally high (worse performance compared to control subjects) in the cued condition. In other words, adults with dyslexia failed to capitalize on the information conveyed by the cue to reduce uncertainty about the target's location. Thresholds in the cued condition distinguished adults with dyslexia from controls better than a variety of other psychophysical and widely used clinical measures (Roach & Hogben, 2007). The difference between good and poor readers was strongest when the set size was largest (16 items). Importantly, the stimuli were not linguistic, which avoided a potential confound in comparing performance between good and poor readers. Overall, Roach & Hogben provided some of the strongest evidence to date that adults with dyslexia have an impairment in covert spatial attention.

One unusual feature of this paradigm is that the peripheral cue could potentially trigger both exogenous and endogenous spatial attention. The cue could be exogenous because it appears peripherally, adjacent to and immediately before the target. It could be endogenous because it always indicates the target's location, so the participant can use that information to voluntarily select the most relevant information. Based on a series of additional experiments, Roach & Hogben (2008) concluded that the primary mechanism of the cueing effect in this paradigm is endogenous. We therefore chose to use this paradigm because the interval between the cue and

the stimuli is short enough to prevent eye movement to the target, a potential pitfall when studying covert attention in young children. We will return to the endogenous vs. exogenous question in the Discussion.

#### 1.4 The present study

We adapted and extended Roach & Hogben's (2007) method to study the development of covert spatial attention and how it differs in children and adults with dyslexia. Although Roach & Hogben focused on set size 16, we used a fixed set size of 8 items in the cued and uncued conditions. With a larger set size, performance may be limited by crowding, which is also known to differ in dyslexia (Callens, Whitney, Tops, & Brysbaert, 2013; Cassim, Talcott, & Moores, 2014; S Gori & Facoetti, 2015; Joo, White, Strodtman, & Yeatman, 2018; Moores, Cassim, & Talcott, 2011). In addition, we increased the size and duration of the peripheral cue and made it red, so it would be salient enough for younger participants. In a subset of adult participants we also replicated a condition with the small black cue used by Roach & Hogben (2004, 2007). Finally, we added a "single stimulus" condition, in which the tilted target is presented alone.

There are advantages to measuring thresholds, which are "stimulus-referred" measurements, rather than reaction time or accuracy with a fixed stimulus. First, we set the stimulus difficulty on each trial with an adaptive staircase that converged on the orientation difference that yields ~75% correct performance. The level of overall task difficulty was therefore equalized across participants, regardless of age or reading skill. Second, we fit psychometric functions with separate parameters for the threshold and the upper asymptote. The latter parameter can capture differences the participants' abilities to stay engaged and follow instructions. Third, we operationalized the effect of attention as the difference in log thresholds between cued and uncued conditions, which can be theoretically linked to a difference in the noise of the internal representations used to make the perceptual decision (see the **Supplemental Material**). That helps us interpret the interactions between cue condition and age or reading ability. Finally, thresholds in the single stimulus condition provide a baseline measure of group differences in the ability to make fine visual discriminations, independent of attention.



## 2. Materials and Methods

We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

### *2.1 Participants*

All participants had normal or corrected-to-normal vision and gave written informed consent in accordance with the University of Washington Institutional Review Board. Flyers solicited participants with and without dyslexia, on campus and through local organizations that provide support to people with dyslexia, including the University of Washington Disability Resources for Students (DRS), Disabilities, Opportunities Internetworking and Technology (DO-IT), and Dyslexic Advantage (<http://www.dyslexicadvantage.org/>). Compensation was a fixed monetary payment.

We recruited a total of 131 participants (68 male) between the ages of 5 and 70 years. The sample size was guided by Roach & Hogben (2007), who used a similar design. We used their data to simulate an ROC analysis of cued thresholds with set size 8, and estimated the sample size required for an 80% probability of detecting a significant difference between the dyslexia group and the control group. That minimal sample size was 12 for each group. Seeking at least that many participants in each group, divided by reading ability and age bin (below and above 20 years), we recruited as many participants as possible in the time window available for the study. Of the 131 participants recruited, two chose to discontinue after only a few practice blocks. Three additional participants were excluded because they reported having an uncorrected vision problem, including amblyopia (criterion determined in advance). Of the remaining 126 participants, 3 children (ages 7-8 years) were excluded because in all conditions of the experiment their accuracy was not significantly above chance (criterion determined after data analysis). That suggested they were not engaged in the task or following instructions (more details in the Analysis section below). The final sample included 123 participants (64

male). Of those, 75 reported having a diagnosis of dyslexia, and 22 reported a diagnosis of attention-deficit/hyperactivity disorder (ADHD), which commonly co-occurs with dyslexia.

Each participant also completed a battery of psychometric tests. Following Roach & Hogben (2007), our primary measure of reading ability was the TOWRE-II Phonemic Decoding Efficiency (PDE) test, which requires speeded reading of pronounceable pseudowords (Torgesen, Rashotte, & Wagner, 1999). Participants also completed the TOWRE-II Sight Word Efficiency sub-test (SWE; speeded reading of real words), and the Wechsler Abbreviated Scale of Intelligence (WASI-III).

To examine the relationship between attention and reading skills, we divided our sample into two groups using a definition consistent with previous work (Roach & Hogben, 2007): individuals with dyslexia (DYS), and individuals with typical reading skills (control participants; CON). Participants in the DYS group (N=46) had a TOWRE PDE score  $\leq 85$  and either reported a history of reading difficulty or a diagnosis of dyslexia. A PDE score of 85 is 1 standard deviation below the age-adjusted norm of 100. Participants in the CON group (N=37) had a TOWRE PDE score  $> 85$  and no diagnosis of dyslexia. This left out a third set of participants (“Neither”, N=40), all but one of whom were diagnosed with dyslexia but a PDE score in the normal range. Demographics and psychonomic data are reported in **Table 1**.

The DYS group differed significantly from the CON group in both total IQ (WASI Full-Scale 2:  $t(78) = 3.66$ ,  $p=0.0005$ ) and non-verbal IQ (WASI Matrix Reasoning:  $t(78) = 2.35$ ,  $p=0.021$ ). Given that the full-scale IQ assessment includes a verbal component, it is not surprising that individuals with dyslexia scored lower. To be certain that any differences in our task performance were not confounded by non-verbal IQ differences, we included the Matrix Reasoning score as a covariate in our analyses of group differences. We use phonemic decoding (TOWRE PDE) as our primary measure of reading ability, to be consistent with a prior study (Roach & Hogben, 2007). In the **Supplemental Material** we present analogous results using the sight word efficiency (SWE) score instead.

Group	Total N	N males	N ADHD Diagnosis	N Dyslexia Diagnosis	WASI Matrix Reasoning	WASI Full-scale 2	TOWRE PDE	TOWRE SWE
DYS	46	27	7	36	50.0 (12.3)	102.4 (15.8)	75.7 (6.8)	78.6 (11.5)
CON	37	19	5	0	55.9 (9.7)	114.6 (13.6)	104.4 (11.3)	101.5 (12.1)
Neither	40	18	10	39	55.4 (8.1)	112.9 (12.1)	95.4 (8.7)	92.6 (12.1)

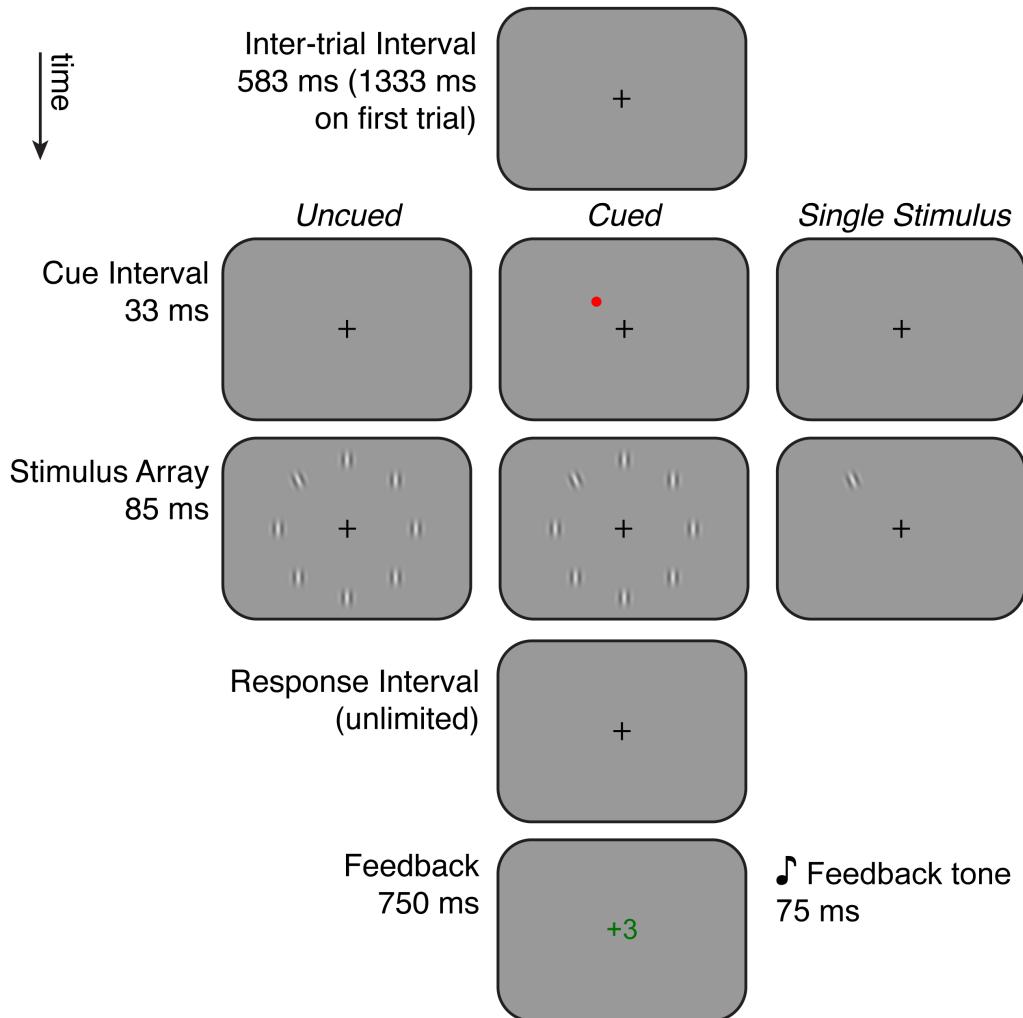
**Table 1.** Demographics of our sample. For the standardized test scores in the last four columns, we report the mean and (standard deviation). The WASI Matrix Reasoning (norm of 50) is our non-verbal IQ measure, and the Full-Scale 2 (norm of 100) score combines verbal and non-verbal components. The TOWRE scores are both normed to 100. PDE = Phonemic Decoding Efficiency (reading aloud pseudowords); SWE = Sight Word Efficiency (reading aloud real words).

A subset of the data from adult participants has been reported in a previous publication (Joo et al., 2018).

## 2.2 Equipment and Stimuli

We generated stimuli using MATLAB (The Mathworks Corporation, Natick, MA, USA) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) on a Linux PC (Mint Mate, version 17). We used an LG liquid crystal display (1920 × 1080 resolution, 120 Hz refresh rate) that subtended 51 degrees of visual angle (°) horizontally. The participant sat at a chinrest 53 cm from the monitor, and used two keys (the down arrow and the right arrow) to make the response. The keys were re-labeled with stickers to indicate the leftmost key was for reporting targets tilted counterclockwise of vertical, and the rightmost key for clockwise of vertical. The surrounding keys on the keyboard were covered with cardboard.

The screen background was set to medium gray. A black fixation cross (0.3°×0.3°) was present at all times except when feedback was given (see below). The target stimuli (**Figure 1**) were Gabor patches: 50% contrast sinusoidal gratings (2 cycles/°) windowed by a 2D Gaussian envelope (SD = 0.28°).



**Figure 1: Stimuli and trial sequence.** The inter-trial, response, and feedback intervals were identical in the 4 conditions, which differed only in the cue and stimulus intervals. Note that this diagram is not to scale. Not shown is the Small Cue condition, in which the cue was a smaller black dot presented for 17 ms immediately preceding the stimulus array.

There were 3 conditions of the experiment: a Single Stimulus condition, and two conditions with eight stimuli: Uncued and Cued. The Cue was a bright red (full saturation) disk  $0.6^\circ$  in diameter presented at the same polar angle as the target,  $3^\circ$  from fixation ( $2^\circ$  from the center of the target).

Participants over the age of 14 were also tested in a condition with a smaller cue designed to match Roach & Hogben (2007). The small cue was a black disk  $0.18^\circ$  in diameter, presented

3.88° from fixation (1.12° from the target) for only 17 ms. Results using this smaller cue were largely consistent as when cued with the bright red cue and are presented in the **Supplemental Material**.

### 2.3 Trial sequence and procedure:

All trials began with an inter-trial interval, during which a fixation cross was displayed on a gray background (**Figure 1**). On the first trial of each block, this fixation interval lasted 1333 ms; in all subsequent trials it lasted 583 ms. Next was a 33 ms cue interval which was blank for the Uncued and Single Stimulus conditions and contained the large red cue in the Cued condition. Immediately after the cue interval, the stimulus array was presented for 83 ms. In all conditions except the Single Stimulus condition (see below), 8 Gabors were presented simultaneously, arranged equidistant from each other in a circle around fixation with radius 5°. Seven of the Gabors were oriented vertically (the *distractors*); one Gabor (the *target*) was tilted either clockwise or counterclockwise (50% probability).

The degree of target tilt was controlled by a staircase procedure (see below). The position of the target in the array was chosen randomly on each trial. In the Single Stimulus condition, the target was presented at a random one of the 8 positions. with no distractors. After the stimulus array, the observer reported the target's tilt direction by pressing one of the two response keys. Response time was unlimited, and accuracy was emphasized. Immediately after the response, a 75 ms feedback tone was presented: a high-pitched (600 Hz) or low-pitched (180 Hz) tone for correct or incorrect responses, respectively. At the same time, visual feedback appeared at the screen center, in the form of the number of "points" won: if the response was correct, "+3" was displayed in green text; otherwise, "+0" in red text. The visual feedback was presented in 25 pt Arial font for 750 ms, after which the next inter-trial interval began.

The participant was instructed that the goal was to win as many points as possible, and that a total of 700 points was needed to "win the game." At the end of each block, the total number of points won that block was displayed, as well as the total number of points gained so far in the session, along with the motivational phrase: "Great job! How many more can you get?"

Trials were conducted in blocks of 52<sup>†</sup>. During 48 of those trials, the degree of the target's tilt was controlled by a weighted 1-up/1-down staircase procedure controlled by the Palamedes toolbox (Prins & Kingdom, 2009). The step size up was set to  $\log_{10}(4^\circ)$  at the start of each staircase, and then halved after four staircase reversals. The step size down was always one third of the step size up, which makes the staircase converge on the 75% correct threshold. The tilt was free to vary between 0.1° and 25°. On a random 4 trials in each block, the target's tilt was set to twice the current threshold estimate and the staircase was not updated. We included these "easy" trials to improve our estimate of the upper asymptote of the psychometric function and provide some relief to untrained observers who may find it difficult to perform at threshold for an extended period of time.

The condition was constant for each block of trials, and was announced before the block began with text on the screen: for Uncued, "There will be no dot, so you have to find the tilted stripes on your own."; for Cued, "The red dot will appear near the tilted stripes."; for Small Cue, "The small black dot will appear near the tilted stripes."; for Single Stimulus, "There will only be one patch of stripes at a time."

Over the course of data collection, we used three slightly different versions of the experiment that differ primarily in which conditions were included. All versions included the Cued and Uncued conditions. Version 2 was used in some of youngest participants and excluded the Single Stimulus condition. All participants 14 or older were tested in Version 3, which included the Small Cue condition. See the **Supplemental Material** for more detail.

Each participant was first trained in the task by viewing pictures of the stimuli. The task was introduced as a "game," with the objective to find a tilted patch of stripes and report which way it is tilted. They practiced the single stimulus condition with easy tilt levels (8 trials per practice block). Then they practiced the uncued condition with increasingly difficult tilt levels until they could perform above chance with less than 15° tilt. Finally, they were introduced to the Cued

---

<sup>†</sup> In Version 2 of the experiment, there was an additional "easy" trial that brought the total number of trials per block to 53 trials. Also in Version 2, the target's tilt on easy trials was set to a fixed value of 25°, rather than twice the current threshold estimate.

condition. The instructions about the cue said, “In the next part of the game, we help you out and show a red dot next to the one patch of stripes that will be tilted. Your job will be to find that red dot, see the stripes that are tilted, and press the correct button. It's important to note that in this game the red dot will ALWAYS be next to the stripes that are tilted and will appear just a split-second before the stripe patches.” For participants in Version 3, similar instructions followed about the Small Cue. Each session lasted roughly 25 minutes. No part of the study procedures or analyses was pre-registered prior to the research being conducted.

## 2.4 Analysis

Our primary measure was tilt thresholds: that is, the degree of tilt required for the participant to discriminate the direction of the target's tilt with 75% correct accuracy, in each condition. To estimate thresholds we used a maximum likelihood procedure in the Palamedes toolbox (Prins & Kingdom, 2009) to fit the raw data in each condition with a log version of the Weibull function (also known as the Gumbel):

$$F(x) = \gamma + (1 - \gamma - \lambda)(1 - e^{-10^{\beta(x-\alpha)}})$$

where  $F(x)$  is proportion correct,  $x$  is the tilt level in  $\log_{10}$ (degrees),  $\gamma$  is chance level (0.5),  $\lambda$  specifies the lapse rate (1 – the upper asymptote),  $\beta$  determines the slope of the curve, and  $\alpha$  determines the location of the function along the  $x$ -axis. Fitting occurred in three stages: We first allowed the three parameters ( $\beta$ ,  $\lambda$  and  $\alpha$ ) to vary freely, with the exception that  $\lambda$  was capped at 0.125. (If  $\lambda = 0.125$ , the subject guesses on 25% of trials, with  $p(\text{correct}) = 0.5$ ). We then fixed the slope  $\beta$  to 1.77, which was the median of freely estimated slopes averaged across subjects and conditions (excluding poor fits with thresholds  $>45^\circ$ ) and refit the remaining two parameters,  $\lambda$  and  $\alpha$ . Finally, we fit again by allowing only the parameter  $\alpha$  to vary, with the parameter  $\beta$  still fixed to 1.77 and the lapse rate parameter  $\lambda$  fixed to the average of  $\lambda$  estimates across all conditions for that subject. Therefore, for each subject, only the parameter  $\alpha$  varied across conditions, but  $\lambda$  could vary across subjects. After fitting, we computed the threshold as the tilt level  $t$  where the  $F(t) = 0.75$ . This procedure avoided over-fitting the data and improved reliability of threshold estimates.

We computed the effect of the cue on thresholds (the “cueing effect”) as the difference in  $t$  between the Uncued and Cued conditions, where  $t$  is in  $\log_{10}$  units. This cueing effect indexes the observer’s ability to capitalize on the cue to process the target and filter out distractors. See the **Supplemental Material** for a computational model that justifies using the difference of log thresholds to measure the attention effect.

Finally, we estimated whether each participant was performing significantly above chance in each condition. For each condition, we estimated the 95% confidence interval of the binomial probability of a correct response. If that confidence interval included 0.5, then we concluded that the participant was not engaged in the task during those trials, and excluded that threshold from the analysis (but included other data from that subject). A total of 9 thresholds from 6 participants (ages 7-14) were excluded for this reason (3 Uncued, 5 Cued, and 1 Small Cue). The median of those excluded thresholds was 140°, and several were effectively infinite.

## 2.5 Reliability

78 participants completed a second session identical to the first, allowing us to compute test-retest reliability by correlating thresholds across the two sessions. With this smaller sample, the average reliability across conditions was  $r = 0.73$ . For the Uncued and Cued conditions (conditions for which we have the most data),  $r=0.75$  and  $0.82$ , respectively. For the cueing effect,  $r=0.46$ .

## **3. Results**

### 3.1 Development of thresholds and spatial cueing effects

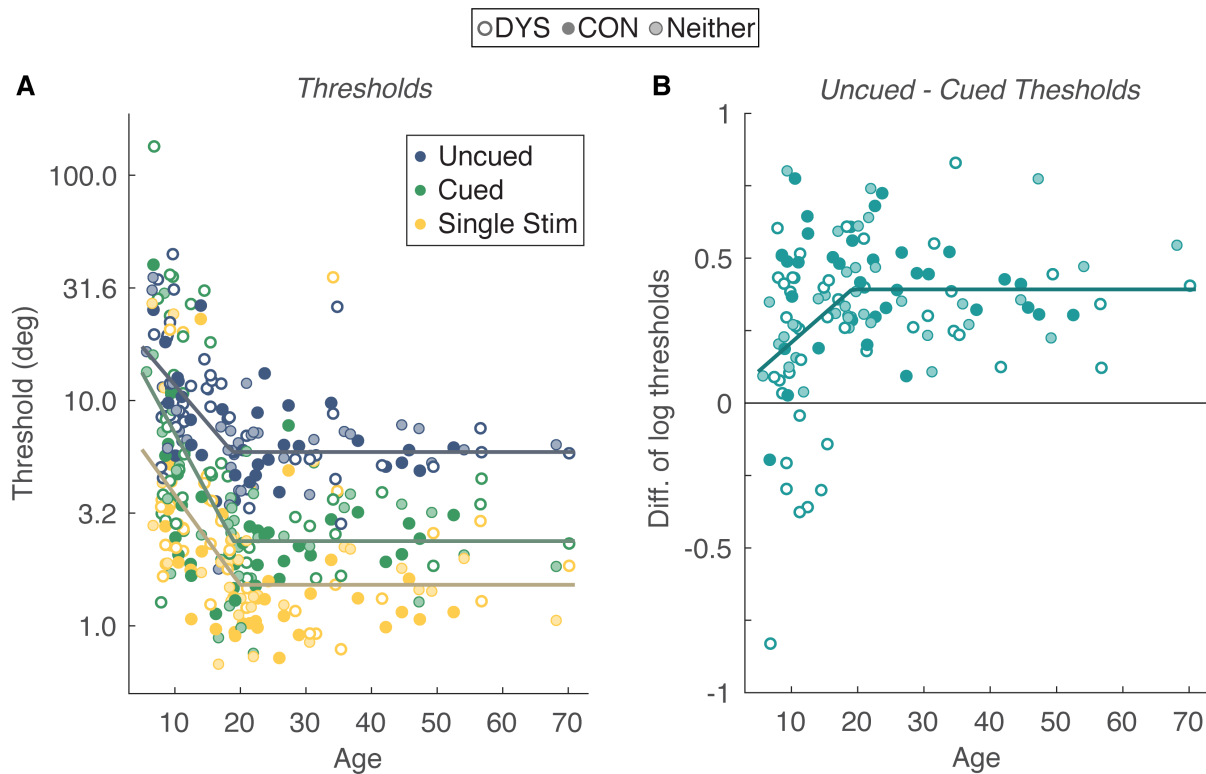
Before analyzing the effect of reading ability on task performance, we model the developmental time-course of visual spatial attention skills. We can then use this model to control for age in the analysis of reading ability. **Figure 2A** shows individual subject thresholds as a function of age in three main conditions: Uncued, Cued, and Single Stimulus. Lower thresholds imply better orientation discrimination. Overall, thresholds start out high in childhood, decrease through adolescence, and then plateau. We fit each condition with a piecewise linear function that



assumes that thresholds ( $t$ ) are a linear function of age ( $A$ ) with slope  $s$  and y-intercept  $b$ , up until an inflection point  $c$  after which there is no more change with age. Specifically:

$$t = sA + b, \text{ if } y \leq c$$

$$= sc + b, \text{ otherwise}$$



**Figure 2: Development of orientation discrimination thresholds and covert spatial attention. (A)** Thresholds as a function of age in each of three conditions. Each point is one threshold from one participant. The fill color of each point indicates the participant’s reading ability group, as indicated by the legend at top. The lines are the best-fitting piecewise linear functions to all subjects **(B)** The difference in log thresholds between the Uncued and Cued conditions as a function of age. Each dot is one participant’s difference. The line is the best-fitting piecewise linear function to all subjects.

We used a least-squares cost function and a leave-one-out cross-validation to assess model fit. The piecewise model fit the data significantly better than a simple linear regression model (constant slope across the age range). For the simple linear model, cross-validated  $R^2$  in each condition (Uncued, Cued, and Single stimulus) was: 0.132, 0.118, and 0.070, respectively.  $R^2$  values were higher for the piecewise linear model: 0.382, 0.307, and 0.217. In addition to the

superior fit to the data, the piecewise model provides an estimate of the age at which performance matures.

The least-squares fits are plotted as solid lines in **Figure 2A**, and the corresponding values of  $s$ ,  $b$ , and  $c$  are listed in **Table 2** along with 95% bootstrapped confidence intervals (CIs). Each CI is the 2.5th and 97.5th percentile on the bootstrapped distribution of parameter estimates (5000 independent samples with replacement). In all three conditions, the initial slope was significantly negative: the 95% CIs excluded 0. The inflection point  $c$  for all 3 conditions was between 18 and 20 years.

<b>Condition</b>	<b>Slope <math>s</math> <math>\log_{10}(\text{deg})/\text{year}</math></b>	<b>Intercept <math>b</math> <math>\log_{10}(\text{deg})</math></b>	<b>Inflection <math>c</math> (years)</b>
<i>Uncued</i>	-0.034 [-0.077 -0.023]	1.41 [1.25 1.81]	18.8 [13.4 20.4]
<i>Cued</i>	-0.054 [-0.170 -0.036]	1.39 [1.11 2450]	18.9 [12.5 20.6]
<i>Single Stim.</i>	-0.040 [-0.075 -0.025]	0.98 [0.72 1.39]	19.9 [16.3 22.8]

**Table 2:** Parameter estimates for the developmental fits to thresholds in each condition. For each parameter, the table lists the best-fitting value and in brackets the 95% confidence interval from bootstrapping.

The fitted functions reveal large differences in thresholds between the three conditions at the asymptotic level (that is, in adulthood). The 95% CIs were all non-overlapping for all three conditions. The large increase in thresholds from the Single Stimulus condition (set size 1) to the Uncued condition (set size 8) reflects the effect of adding the 7 vertical distractors to the target display. The distractors add noise to the orientation estimation process and may tax limited processing resources. The reduction in thresholds from the Uncued to the Cued condition is the effect of covert spatial attention. The mean cued threshold was 59% of the way between the Uncued threshold and the lower bound represented by the Single Stimulus threshold. Thus, the cue allows the typical adult to filter out much of the noise added by the distractors, but not all of it (Roach & Hogben, 2007).

To compare the rate of developmental change across conditions, we computed 95% bootstrapped CIs on the difference in slope parameters. The slope of developmental change in the Cued condition was 59% greater than in the Uncued condition, and the 95% CI of differences excluded zero: [0.001 0.119]. The slope in the Cued condition was 35% greater than in the Single Stimulus condition, but the CI included zero: [-0.012 0.15]. The slopes in the Uncued and Single Stimulus conditions did not significantly differ: CI = [-0.020 0.030].

To directly assess the development of spatial attention we analyzed the difference in log thresholds between the Uncued and Cued conditions (i.e., the cueing effect). **Figure 2B** plots the cueing effect for each participant as a function of age. (Note that this analysis includes 117 participants, excluding 6 who did not perform above chance in either the Cued or Uncued conditions). We fit those data with the same piecewise linear function as in Figure 2A. The best-fitting slope was  $0.020 \log_{10}(\text{deg})/\text{year}$  (95% CI = [0.008 0.166]), indicating a significant developmental increase in the cueing effect, above and beyond the change observed in each individual condition. The inflection point was 19.2 years (CI = [10.2 22.8]).

In sum, the developmental analysis of orientation discrimination thresholds revealed two patterns: (1) A large improvement in orientation discrimination ability, as measured in the Single Stimulus condition, from age 5 to about 20. The best-fitting developmental function predicts a threshold of  $6.1^\circ$  at age 5, and a threshold of  $1.5^\circ$  after age 20. (2) A difference in thresholds between the Cued and Uncued conditions that also increases with age up until about 20 years. Therefore, in addition to the overall improvement in visual sensitivity with age, there is also an improvement in the attentional selection of information from relevant peripheral visual field locations. However, as shown in the next section, the developmental increase in the cueing effect was primarily driven by participants with poor reading ability.

### 3.2 Comparison of spatial attention in good and poor readers

#### *3.2.1 Thresholds in individual conditions*

We next evaluated the association between reading ability and orientation discrimination thresholds, and between reading ability and the cueing effect. We took two complementary

approaches: (1) comparing two sub-groups of our sample, those with and without dyslexia; (2) correlating the TOWRE PDE score with threshold measurements in the whole sample. The first group comparison approach is typical in the literature and was used in the prior study of adults that reported group differences in this task (Roach & Hogben, 2007). Here we examine whether those differences between good and poor readers are also present in children, before performance in this task matures. The second approach treats reading ability as a continuum (Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992) and includes all participants.

Following the first approach, we defined two sub-groups of our sample: The “DYS” group included all individuals who had a PDE score  $\leq 85$  and either reported a history of reading difficulty or a diagnosis of dyslexia. The control group, “CON”, included all individuals with a PDE score  $> 85$  and no diagnosis of dyslexia. This left out roughly 1/3 of our participants, who indicated a diagnosis of dyslexia but had a PDE score in the normal range. We further divided the DYS and CON groups into two age bins: 5-19 years and 20-70: based on the developmental patterns described above, the first age bin is in the range when orientation discrimination thresholds as well as cueing effects are still maturing. **Table 3** reports the number of subjects in each age bin and reading ability group.

Ages	DYS	CON	Neither
5-19	31	18	24
20-70	15	19	16

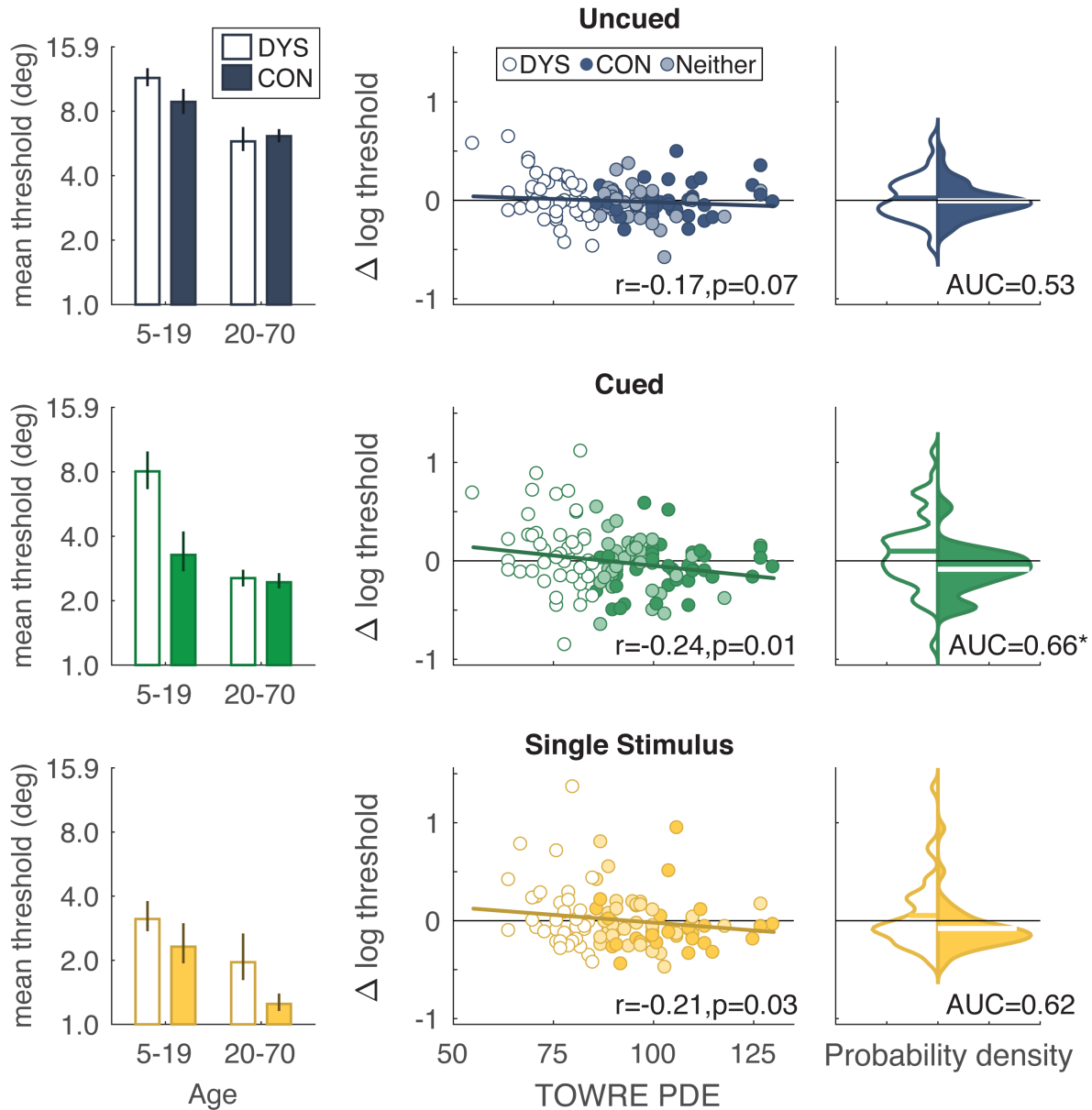
**Table 3:** Counts of subjects in each of two age bins and three groups sorted by reading difficulty (DYS = individuals with dyslexia; CON = control subjects with typical reading ability).

The mean thresholds in each group are plotted in **Figure 3, left column**. In the younger age bin, thresholds are higher in the DYS group than the CON group, especially in the Cued condition. To analyze these differences in thresholds in the younger age bin we fit a LME model with fixed effects of condition and reading ability group, and a random effect of subject. We also included fixed effects of age, non-verbal IQ (matrix reasoning score), and ADHD diagnosis as covariates. Consistent with the developmental patterns reported above, the effect of age was significant

( $F(1,114)=12.3$ ,  $p=0.0007$ ). ADHD diagnosis had no significant effects or interactions ( $p>0.40$ ). Higher non-verbal IQ was associated with lower thresholds overall ( $F(1,114)=3.59$ ,  $p=0.061$ ), but that did not interact with condition ( $F(2,114)=1.09$ ,  $p=0.34$ ). There was a large main effect of condition on thresholds ( $F(2,114)=32.8$ ,  $p<10^{-11}$ ). On average, thresholds in the DYS group were 0.13  $\log_{10}(\text{deg})$  higher than in the CON group, but the main effect of reading ability group was not significant ( $F(1,114)=2.38$ ,  $p=0.126$ ). However, there was an interaction between condition and reading group ( $F(2,114)=3.56$ ,  $p=0.032$ ). As seen in the left column of Figure 3, the relative elevation of thresholds DYS group under 20 years old was greatest in the Cued condition.

The same analyses of thresholds within the adult age bin (over age 20) found no effect of age ( $F<1$ ), consistent with the conclusion above that thresholds asymptote after age 20. ADHD diagnosis had no effect ( $F<1$ ), but there was a main effect of non-verbal IQ ( $F(1,78)=18.5$ ,  $p<10^{-4}$ ). Thresholds differed significantly across conditions ( $F(2,78)=36.1$ ,  $p<10^{-11}$ ), but there was no main effect of reading ability group ( $F(1,78)=0.36$ ,  $p=0.55$ ; mean difference = 0.03). There was, however, an interaction between condition and reading group ( $F(2,78)=4.15$ ,  $p=0.019$ ).

The bar plots in the left column of Figure 3 suggest that the threshold difference between the DYS and CON groups is larger in participants below age 20 than in participants above age 20, at least for the Cued and Uncued conditions. Children with dyslexia may become more similar to control participants as they reach adulthood, in terms of visual search performance. Such a hypothesis predicts an interaction between age bin and reading ability group on thresholds. Controlling for ADHD diagnosis, non-verbal IQ, and age within each group, that interaction was significant in the Cued condition ( $F(1,68) = 4.34$ ,  $p=0.041$ ). The interaction was not significant in either the Uncued condition ( $F(1,70)=2.12$ ,  $p=0.15$ ) or the Single Stimulus condition ( $F(1,57)=0.33$ ,  $p=0.57$ ). To gain more power, we conducted a similar analysis but including all participants and using the TOWRE PDE score as a continuous measure of reading ability. The interaction between reading ability and age bin was significant for both the Uncued ( $F(1,106)=5.50$ ,  $p=0.021$ ) and Cued ( $F(1,105)=7.77$ ,  $p=0.006$ ) conditions, but not the Single Stimulus ( $F(1,91)=0.90$ ,  $p=0.34$ ). Therefore, for both conditions with set size 8, the relative impairment in poor readers becomes less severe after age 20.



**Figure 3.** Threshold differences between good and poor readers. **Left column:** mean thresholds for subjects divided into two age groups (left pair of bars: 5-19 years; right pair of bars: 20-70 years) and two reading groups (DYS and CON). Error bars are bootstrapped 68% CIs. **Middle column:** correlations between TOWRE phonemic decoding score (PDE) and the residuals of the developmental model plotted in Figure 2A. This includes *all* subjects. The y-axis is the difference between each subject's threshold and the threshold predicted by the developmental model. The solid line is the prediction based on PDE score from a model that also included ADHD and non-verbal IQ as predictors. **Right column:** smoothed probability density functions of the residuals plotted in the middle column, just for the DYS and CON groups. The Gaussian smoothing kernel bandwidth was set to 0.06. The horizontal line superimposed on each distribution is the mean. The AUC is the area under the ROC curve. The asterisk indicates significant difference from 0.5.

In the next analysis, we ask how reading ability can explain the residual variance in thresholds (and cueing effects) that was not explained by age in the developmental model shown in Figure 2. This controls for the effect of age while maximizing statistical power to estimate the effect of reading ability by using the entire sample. In the middle column of **Figure 3** we plot the relation between each subject's reading score and the *residuals* from the piecewise linear models of thresholds as a function of age for each condition. The residuals are the differences between each subject's threshold and the prediction of the developmental model. Each subject's reading group (DYS, CON or Neither) is indicated by the color of the dot. In all three conditions (Uncued, Cued, and Single Stimulus), there was a significant negative correlation between reading score and threshold residuals. This means that better readers are likely to have lower thresholds compared to poor readers of the same age. Therefore, not all of the unexplained variance by the developmental model in Figure 2A is just noise; some of it is explained by reading ability. We further analyzed these residuals with a linear model that included reading score (TOWRE PDE), ADHD diagnosis, and non-verbal IQ. ADHD had no effects in any condition, but higher non-verbal IQ was associated with lower thresholds in the Uncued (slope=-0.005;  $t(108)=2.50$ ,  $p=0.014$ ) and Single Stimulus (slope=-0.009;  $t(91)=3.15$ ,  $p=0.002$ ) conditions. The effect of reading score was significant in the Cued condition (slope=-0.004;  $t(109)=2.02$ ,  $p=0.046$ ), but not in the Uncued (slope=-0.001,  $t(110)=1.05$ ,  $p=0.30$ ) or and Single Stimulus conditions (slope=-0.003,  $t(95)=1.58$ ,  $p=0.12$ ). The solid lines in the middle column of Figure 3 are the predictions based on reading score alone.

We conducted another analysis to examine how well these psychophysical measures distinguish dyslexic from control participants, as suggested by Roach & Hogben (2007). Specifically, we conducted a Receiver Operating Characteristic (ROC) analysis of the distributions of threshold residuals. The area under the ROC curve (AUC) is a measure of how accurately we can classify subjects into the *DYS* or *CON* groups, ranging from 0.5 (chance) to 1.0 (perfect).

In the right column of **Figure 3**, we plot smoothed distributions of residuals (same data as in the middle column) for the two groups (*DYS* in white, *CON* in the darker color). To evaluate

whether each AUC was significantly higher than predicted by chance, we performed a permutation test to establish a null distribution: on each of 5000 repetitions, we randomly shuffled the group labels for all subjects and re-computed the AUC. The actual AUC exceeded the 95% CI on this null distribution in the Cued condition: AUC = 0.66 (null CI = [0.37 0.63]). As seen in Figure 3, the two groups' distributions in the Cued condition greatly overlapped, with a rather subtle upward shift for the DYS group (mean =  $0.10 \pm 0.06$ ) compared to the CON group ( $-0.08 \pm 0.04$ ). AUC was not significantly above chance in Uncued (0.53, null CI = [0.37 0.63]) or Single Stimulus conditions (0.62, [0.36 0.64]).

In summary, poor readers have higher orientation discrimination thresholds than good readers, especially in the Cued condition. Moreover, in the Cued condition, the difference between good and poor readers is greater before age 20 than in adulthood. That interaction between reading ability and age is less robust in the Uncued condition and appears to be absent in the Single Stimulus condition.

### 3.2.2 *The cueing effect on thresholds*

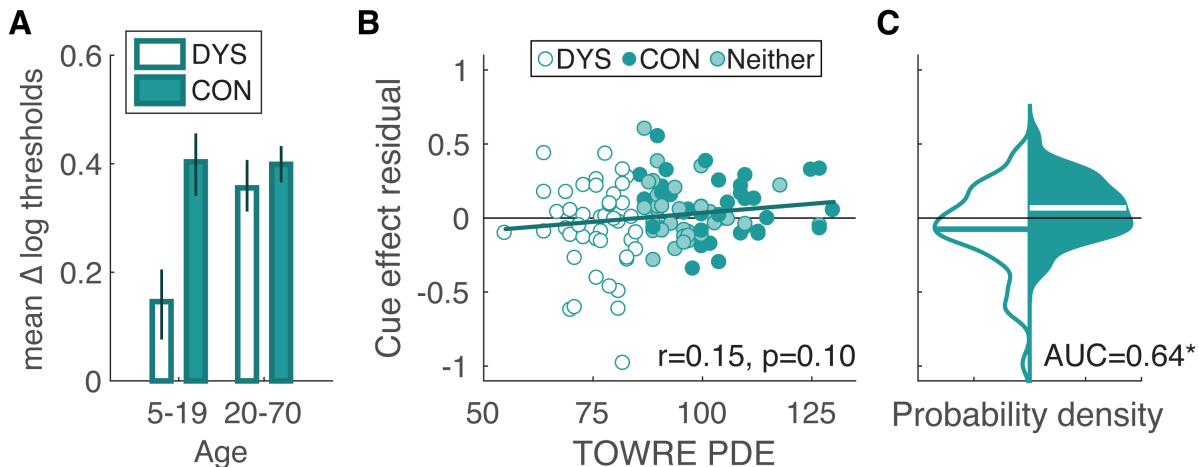
To specifically evaluate spatial attention, we analyze the relation between reading ability and the *cueing effect*. **Figure 4A** plots the mean cueing effects for the DYS and CON groups, for those under and above 20 years of age (as in Figure 3). For each age bin, we analyzed the cueing effect with a linear model that included reading group, age, ADHD diagnosis, and non-verbal IQ as predictors. In the younger age bin, the model estimated that the cueing effect is 0.22  $\log_{10}(\text{deg})$  greater in the CON group than in the DYS group ( $t(41) = 2.17, p=0.036$ ). The effect of age was not significant ( $t(41)=1.67, p=0.10$ ), and there were no effects of ADHD diagnosis or non-verbal IQ (both  $t < 1$ ). Therefore, individuals under 20 with dyslexia have (on average) weaker attentional selection than those without dyslexia.

In the adult age bin (over 20), a similar linear model estimated that the cueing effect was 0.06  $\log_{10}(\text{deg})$  larger in the DYS group than the CON groups, but was not statistically reliable ( $t(26) = 1.04, p=0.31$ ). The effect of age ( $t(26)=1.50, p=0.15$ ) and non-verbal IQ ( $t(26)=1.86, p=0.074$ ) also were weak. There was a reliable effect of ADHD diagnosis: the model estimated that adults



with ADHD have a cueing effect  $0.18 \log_{10}(\text{deg})$  smaller than adults without ADHD ( $t(26)=2.15$ ,  $p=0.041$ ). However, note that this analysis only included 4 adults with ADHD.

Altogether, the data in Figure 4A show that children with dyslexia have a markedly smaller spatial cueing effect than typically-reading children, but adults with dyslexia have a cueing effect that is only slightly smaller than typical. That suggests that as individuals with dyslexia mature, they recover some spatial attention function; i.e., the differences in childhood reflect a developmental delay rather than a deficit *per se*. We urge caution in interpreting that result, however, because it is not supported by a statistical interaction between reading ability group and age bin. Specifically, we fit a linear model of the Uncued-Cued log threshold difference, with factors age bin (under 20; over 20), reading group (DYS, CON), and their interaction. We also include covariates of the normalized (mean 0) age within each age group, ADHD diagnosis, and non-verbal IQ. The cueing effect was smaller in the DYS group ( $F(1,69)=7.72$ ,  $p=0.007$ ). However, the interaction between age bin and reading group was not significant ( $F(1,69)=2.32$ ,  $p=0.132$ ), nor were any of the other main effects. To maximize power, we conducted a similar analysis but using all participants and using the TOWRE PDE score as a continuous measure (normalized to mean 0 within each age bin), rather than dividing by reading group. Consistent with the developmental pattern reported above, there was a significant effect of age bin ( $F(1,105) = 6.95$ ,  $p=0.010$ ). No other factor had a significant effect except reading score ( $F(1,105)=4.58$ ,  $p=0.035$ ). Again, the effect of reading score did not interact with age bin ( $F(1,105)=2.23$ ,  $p=0.138$ ).



**Figure 4.** Relation between the cueing effect on thresholds and reading ability. **A:** Mean cueing effects, expressed as the difference in log thresholds between the Uncued and Cued conditions. Format as in Figure 3 (left column). **B:** Correlation between reading score and the residuals from the developmental model of the cueing effect. For each subject, the residual is the difference between their measured cueing effect (Uncued – Cued thresholds) and the prediction of the piecewise linear model of age. The solid line is the prediction based on PDE score from a model that also included ADHD and non-verbal IQ as predictors. **C:** Smoothed probability density functions of the residuals plotted in the middle column, just for the DYS and CON groups. The horizontal line superimposed on each distribution is the mean. The asterisk indicates that the AUC differs significantly from 0.5.

Finally, we examined how the residuals of the cueing effect from the developmental model in Figure 2C relates to reading ability. Those residuals, for all participants, are plotted as a function of reading score in **Figure 4B**. The correlation was positive but not significant ( $r=0.15, p=0.10$ ). The solid line is the prediction of a linear model that also included predictors for non-verbal IQ and ADHD. None of those factors were significant, including reading score (slope=0.002;  $t(108)=1.51, p=0.134$ ). However, the ROC analysis yielded a significant difference in residuals between just the DYS and CON groups (**Figure 4C**). The AUC was 0.64 (null CI = [0.37 0.63]). The groups significantly differ even when controlling for ADHD and non-verbal IQ: the mean residual for DYS was  $-0.075 \pm 0.04$ ; and for CON it was  $0.069 \pm 0.03$  ( $t(73) = 2.47, p=0.016$ ). Again, the effect is subtle and the two distributions largely overlap.

In summary, we found that individuals with dyslexia benefit less from the cue than control participants, in line with an impairment of selective covert spatial attention. Previous studies have only reported such differences in selective attention for adults with dyslexia (Roach &

Hogben, 2007). In our data, the difference associated with dyslexia actually appears stronger in people under 20. However, the interaction between age group and reading score on cueing effects was not statistically significant ( $p > 0.05$ ).

### 3.2.3 Analyses in a sliding age window

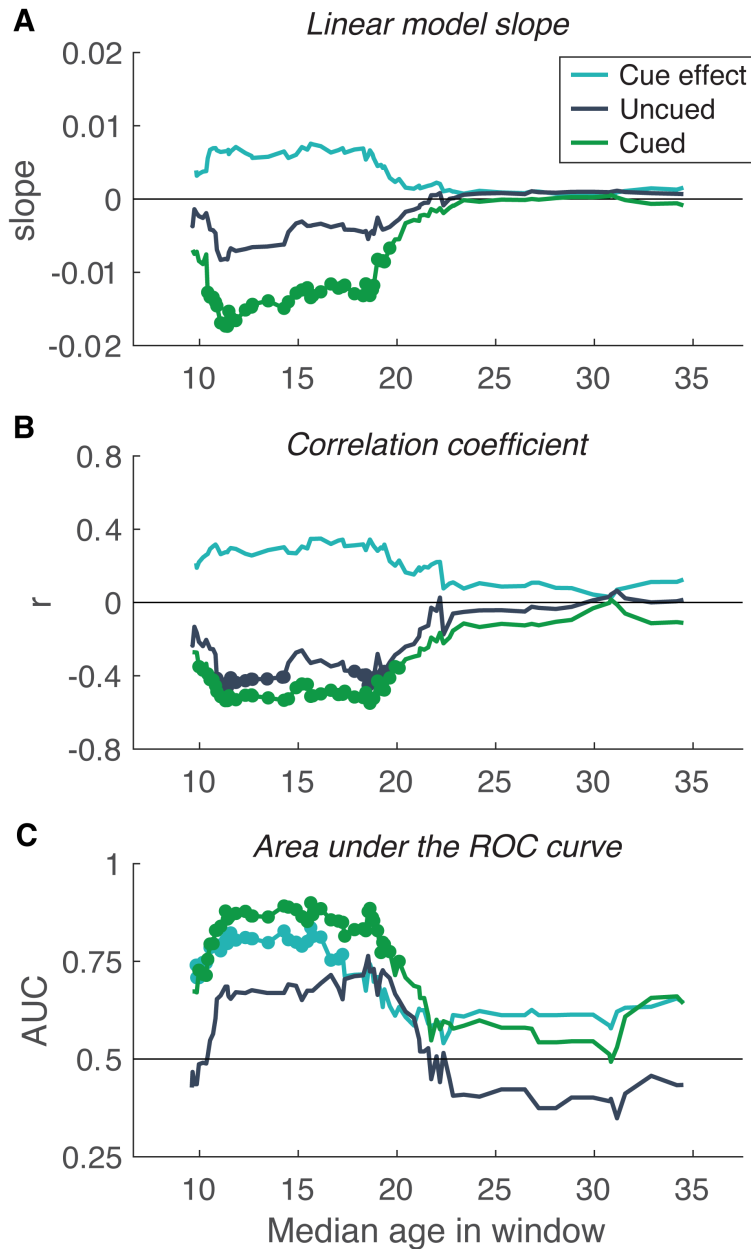
Given that there were significant interactions between reading ability and age bin in some conditions (Figure 3), we sought a finer-grained analysis of developmental changes in the effect of reading score on task performance. **Figure 5** shows the association between reading score and residuals from the piecewise linear developmental models computed in a sliding window across the age range. At each step, the window included 40 participants (1/3 of the sample with usable thresholds). Using the threshold residuals controls for the effect of age *per se*. At each step of the window, we analyzed the Cued and Uncued conditions as well as the cueing effect, with the same three analyses as in the middle and right columns of Figures 3 and 4:

- (1) We fit the residuals with a linear model that had three predictors: TOWRE PDE reading score, ADHD diagnosis, and non-verbal IQ. In **Figure 5A** we plot the estimated slope relating reading score to the residuals, as a function of the median age in the window. Dots indicate points when the slope was significantly different from 0, correcting for multiple comparisons (false discovery rate = 0.05; Benjamini & Hochberg, 1995). There were significantly negative slopes for the Cued condition between median ages 10 and 20 years. Negative slopes mean better readers have lower thresholds (better sensitivity) than predicted by age alone. Within that range, the slope for the Cued condition was roughly twice the slope for the Uncued condition, indicating that good and poor readers differ in how well they can use the cue. After about age 20, the slopes approach 0. The slope relating the cueing *effect* to reading ability starts out positive and then declines after age 20, suggesting that better readers have larger cueing effects in late childhood and adolescence. However, the slopes for cueing effects were not significant at any individual time point.
- (2) We also computed the linear correlation between TOWRE PDE and the residuals in each time window (**Figure 5B**). The correlation coefficients show a similar pattern as the linear

model slopes, but are roughly equivalent for the Cued and Uncued conditions, even before age 20. Therefore, reading ability accounts for the same amount of variance in thresholds for those two conditions, but the difference between good and poor readers is greater in the Cued condition (as reflected by the slopes in Fig. 5A).

- (3) Finally, within each time window we computed the area under the ROC curve that discriminates the DYS and CON groups (**Figure 5C**). Unlike the first two analyses, this takes into account dyslexia diagnosis and history of reading difficulty, and excludes the set of participants who fall into neither group. The Cued condition was significantly above chance between the median ages of 10 and 20, and remained at or above 0.85 between 11 and 19. The Uncued condition was not significantly above chance at any time-point. The cueing effect, however, was significantly above chance between median ages of 10 and 19, reaching a peak value of 0.84 at 16. Therefore, there may be a special link between spatial attention and dyslexia *per se* (Fig. 5C) that is not captured as cleanly by correlating the cueing effect with a single measure of reading ability (the TOWRE PDE score, as used in Figs. 5A and 5B).

This data-driven analysis confirms that the division into two age bins at age 20, as used in Figures 3 and 4, is appropriate. It is notable that the link between reading ability and age-normed task performance diminishes at about the same age as absolute thresholds reach mature levels (Fig. 2). However, note that when we analyzed cueing effects in just two age bins, we did not find a statistically significant interaction between age bin and reading ability group (Fig. 4A).



**Figure 5.** Association between reading ability and task performance assessed in sliding windows across the age range. (A) Slopes of a linear fit between TOWRE PDE scores and *residuals* of the developmental model for Uncued thresholds, Cued thresholds, and the Uncued - Cued effect. The linear model included ADHD diagnosis and non-verbal IQ as covariates. Each time window included 1/3 of the sample (N=40). (B) Similar analysis for the linear correlation coefficients between residuals and TOWRE PDE. (C) Similar analysis of the Area under the ROC curve that distinguishes the DYS group from the CON group. In all three panels, dots on individual time-points indicate that the slope significantly different from 0 (corrected for multiple comparisons).

### 3.3 Testing the effect of cue salience

The ROC analysis above indicated that cued thresholds discriminate the DYS and CON groups. However, the accuracy of that discrimination diminishes in adulthood (Fig. 5C). A prior study with a similar design reported larger group differences in adults than we observed here (with set size 16; Roach & Hogben, 2007). Their cue was a small black dot presented for 20 ms. We increased the salience of the cue for use with children by making it larger and red, with duration 33 ms. These differences motivate an intriguing hypothesis that attentional deficits in dyslexia can be (partially) ameliorated by increasing the salience of the cue. To investigate whether participants with dyslexia are relatively more impaired with a less salient cue, 76 participants over the age of 14 were additionally tested in a Small Cue condition that was matched to Roach and Hogben (2007).

Overall, performance with the Small Cue was similar to performance with the big cue. Indeed, thresholds in those two conditions were highly correlated ( $r=0.68$ ,  $p<10^{-11}$ ). Reading ability predicted thresholds in the Small Cue condition ( $r=-0.25$ ,  $AUC = 0.71$ ), replicating the phenomenon reported by Roach & Hogben (2007). However, the DYS and CON groups differed only slightly in the small cue's *effect* (difference from Uncued). Overall, we found no evidence that individuals with dyslexia perform more differently from good readers when the cue is less salient. See **Figure S1** and text in the **Supplemental Material** for more detail.

### 3.4 Lapse Rates and Response Times

When fitting psychometric functions we also estimated the difference  $\lambda$  between the upper asymptote and 1 (where 1 indicates perfect performance).  $\lambda$  estimates how often the subject's response is uncorrelated with the stimulus, as would result from a lapse of focus.  $\lambda$  was fixed to be equal across conditions but varied across subjects. In our sample,  $\lambda$  decreased with age, suggesting that as children mature they become more consistently engaged in the task and therefore have lower lapse rates. This developmental pattern is in addition to the improvements to orientation sensitivity and selective spatial attention that were apparent in thresholds. Importantly, individuals with dyslexia do not differ from control participants in lapse rates.

Thus, group differences in thresholds cannot be attributed to differences in ability to focus on the task. See the **Supplemental Material** for the full analysis of how  $\lambda$  varies with age and reading ability.

We also analyzed response times (RTs) on correct trials. Although the task was not speeded and not designed to measure attention effects on RTs, it is important to verify that there were no speed-accuracy trade-offs that could explain differences in thresholds between participant groups. For instance, individuals with dyslexia may have less of a cueing effect because they respond too fast in the Cued condition. Overall, there was no relation between reading ability and RT, although among adults over age 20, the DYS group tended to respond slower than the CON group. On average there was no cueing effect on RTs (difference between the Uncued and Cued conditions). Better readers tended to have larger cueing effects, consistent with the pattern in thresholds. Overall, there was no sign of differing speed-accuracy tradeoffs between groups. See the **Supplemental Material** for a full analysis of RTs.

### 3.5 Associations with phonemic decoding vs. real word reading

In the analyses presented above, we used the TOWRE phonemic decoding efficiency score (PDE) to define reading ability groups and to correlate with thresholds (as did prior studies with this paradigm; Roach & Hogben, 2007). The PDE test requires rapid reading aloud of pseudowords. Our participants also completed a similar test with real words (sight word efficiency, SWE). *Sight word reading and phonemic decoding skills, respectively, may distinguish 'surface' and 'phonological' subtypes of dyslexia (McDougall, Borowsky, MacKinnon, & Hymel, 2005), and they may also be differentially related to the neural circuits of exogenous and endogenous spatial attention (Ekstrand, Neudorf, Gould, Mickleborough, & Borowsky, 2019).*

Therefore, we conducted an exploratory analysis of our data using the SWE score in place of the PDE. *Fewer subjects fell into the DYS group based on SWE (N=40) than based on PDE (N=46).* Overall, the results were quite similar. One possible difference is that the association between SWE and cued thresholds does not decrease with age as much as it did for the PDE. More detail,

including the sliding age window analysis, is reported in the **Supplemental Material** (Figure S5). Future studies should explore the hypothesis that adults with a specific impairment in real word reading (rather than just phonemic decoding) have a more robust deficiency in covert spatial attention.

## 4. Discussion

### 4.1 Summary

The primary findings of this study concern the development of orientation discrimination sensitivity and covert spatial attention, and how both of those measures relate to reading ability. First, when analyzing our entire sample together, we found that both orientation sensitivity and spatial attention improve gradually up until around age 20, consistent with previous findings (Brodeur & Enns, 1997; Leclercq & Siéroff, 2013; Schul et al., 2003). Although our sample included only 6 participants over age 50 (aging was not our focus), we observed no decline in task performance or attentional selection in the later years (Folk & Hoyer, 1992; Hartley, Kieley, & Slabach, 1990).

Second, individuals with dyslexia tended to have higher orientation discrimination thresholds and weaker spatial cueing effects than control participants with typical reading skills. Poor reading ability was associated with higher thresholds even in the single stimulus condition, when attentional selection was not necessary (there were no distracting stimuli to filter out). Nonetheless, the difference between groups was largest in the Cued condition, and reading ability correlated with the cueing effect (the benefit relative to the Uncued condition). This pattern shows that many individuals with dyslexia struggled to take advantage of the cue to base their decision on the relevant target stimulus and ignore irrelevant locations. Therefore, dyslexia is associated with a difference in the mechanisms of selective spatial attention.

Surprisingly, the links between reading ability and task performance were strongest in participants below age 20, prior to the maturation of absolute threshold levels (Fig. 5). That means that the attentional advantage in skilled readers is not a *consequence* of decades of fluent



reading experience (Franceschini et al., 2012). That advantage may emerge early during the acquisition of literacy, or be linked to another trait that facilitates both reading and spatial attention from a young age.

#### 4.2 Dyslexia in childhood and in adulthood

Another intriguing hypothesis consistent with our data is that individuals with dyslexia have a developmental delay in visual spatial attention. Eventually their spatial attention approaches normal function, but deficits present in childhood have a lasting impact on their reading ability. That hypothesis requires further investigation, because although we did find a statistically significant interaction between age bin and reading ability for thresholds in the cued condition, that interaction was not quite reliable for the cueing *effect* (relative to the Uncued condition).

Nonetheless, it is worth noting that we found relatively little age-related change in the cueing effect among skilled readers. That is consistent with prior findings that endogenous attention is adult-like by age 10 (Goldberg et al., 2001; Michael et al., 2013; Ristic & Enns, 2015; Wainwright & Bryson, 2005). The more gradual developmental patterns shown in Figure 2 were mostly driven by the poor readers, who were disproportionately represented in our sample. Another limitation is that our study was not longitudinal. The differences observed between children and adults could be due to sampling from different populations. In other words, it is possible that the particular dyslexic 10-year-olds in our sample will continue to have large deficits late into adulthood. Also, the differences between children and adults were somewhat less pronounced when using sight word efficiency to assess reading ability, rather than phonemic decoding efficiency (**Figure S5**).

For those reasons and others, we do not conclude that there is *no* association between spatial attention and reading ability in adulthood. Previous work with the same paradigm found large differences in cued thresholds between adults with and without dyslexia (Roach & Hogben, 2004, 2007, 2008). They varied the set size and found that the cueing effects and group differences were maximal with set size 16. We used set size 8 to reduce the influence of crowding in the periphery, and we were able to roughly replicate their finding that cued

thresholds correlate with reading ability (Fig. S1). But the relatively small effects of reading ability in adults with set size 8 – especially for the uncued-cued threshold *differences* – suggest that the developmental trajectory of reading ability and spatial attention may interact with crowding.

When interpreting the differences between group averages, it is important to note the great deal of overlap in cueing effect magnitudes between individuals with and without dyslexia. Some individuals with dyslexia had large cueing effects, and some excellent readers had no cueing effect. A deficit of covert spatial attention is therefore unlikely to be the single cause of dyslexia. One possible interpretation is that a sub-type of dyslexia is associated with a specific deficit in visual-spatial attention, while many individuals struggle to read for a constellation of other reasons (e.g., crowding; Joo et al., 2018). Another interpretation is that a developmental delay in spatial attention interacts with other deficits to increase the risk that a child will struggle learning to read. Our data highlight the importance of taking a developmental approach. Longitudinal studies combining multiple measures in large samples of children and adults will be needed to resolve these different hypotheses.

#### 4.3 Measurements and theory

We recommend that future studies also measure visual discrimination thresholds. When the primary dependent measure is response time or proportion correct, overall differences in motor function, perceptual ability or cognitive ability can complicate comparisons of attentional effects across groups. The study reported here is the first to measure thresholds in an investigation of spatial attention in children as well as adults with dyslexia. Because we assessed thresholds with an adaptive staircase, overall difficulty was equated across participants. We were also able to separately analyze lapse rates and found little differences between good and poor readers. Therefore, group differences in the cueing effect cannot be explained by differences in overall visual ability or focus on the psychophysical task.

The use of thresholds also allows us to make detailed inferences about the underlying mechanisms driving the effects in each condition. First, consider the age-related improvement

in thresholds in the Single Stimulus condition. The target appeared alone and its contrast was supra-threshold, so performance was not limited by difficulty in target detection or localization. Rather, performance was limited by noise in the visual estimation of its orientation and the formation of a categorical decision (left or right of vertical). Improvement in thresholds with age can be interpreted as a sharpening of orientation-tuned mechanisms early in the visual system (Gilbert, Sigman, & Crist, 2001), or perhaps a more efficient read-out of visual neurons (Petrov, Doshier, & Lu, 2005). Independent of the effect of age, we also found that better readers had lower thresholds in the single stimulus condition (Figure 3, bottom middle panel). Although this effect is small, it suggests that better readers are generally better at discriminating fine visual details, which may not be surprising. There could be many reasons why an individual with dyslexia has elevated orientation discrimination thresholds, but it would be difficult to explain by a pure deficit in the ‘magnocellular’ visual stream (Skottun, 2000).

Second, consider the Uncued condition, when 8 Gabor patches are presented simultaneously. Seven are vertical distractors, and one is the tilted target that appears at an unpredictable location. To perform the task, the observer could either integrate the estimated orientations of all 8 items (by summing or averaging, as suggested by Baldassi & Burr, 2000), or they could report the tilt direction of the one item that appears to deviate most from vertical (search based on a max rule, as suggested by Roach & Hogben, 20007). Either way, the task is difficult because of noise in the estimated orientation of each item. Errors can occur when at least one distractor is mistakenly perceived as being tilted in the direction opposite the target and to a larger degree. Thresholds in the Uncued condition are therefore higher than in the Single Stimulus condition: more signal is required to overcome the noise added by each distractor. Another potential explanation of the threshold elevation in the Uncued condition is that the distractors tax limited processing resources, such that each item’s orientation estimate is noisier than in the Single Stimulus condition.

Third, consider the Cued condition and how the presence of the informative cue affects thresholds. The cue reveals the target’s location just prior to the onset of the eight Gabors. Cued thresholds are far below Uncued thresholds, but do not reach Single Stimulus levels. The benefit

compared to the Uncued condition is the effect of selective spatial attention, and there are several potential explanations for it. One is *signal enhancement*: the cue increases the precision of the perceptual representation of the target (Cameron, Tai, & Carrasco, 2002; Cameron, Tai, Eckstein, & Carrasco, 2004; Carrasco & Yeshurun, 1998; Lu & Doshier, 1998). Another possibility is *uncertainty reduction* (or distractor exclusion; Morgan, Ward, & Castet, 1998; Palmer, 1994; Palmer, Ames, & Lindsey, 1993). Under this hypothesis, the cue allows the observer to base their decision on information at the target location and exclude the noise added by the distractors. Thus, the cue benefits performance even without changing the quality of the underlying perceptual representations. Roach & Hogben (2007, 2008) concluded that the cueing effect in this task is primarily due to uncertainty reduction, although that does not preclude some signal enhancement as well. In the **Supplemental Material** we describe in more detail a quantitative model that links cueing effects to differences in log thresholds.

#### 4.4 Exogenous vs. endogenous spatial attention

To further understand the cueing effect we must classify it as *exogenous* (stimulus-driven, involuntary, and transient), *endogenous* (goal-driven, voluntary and sustained; Carrasco, 2011), or perhaps both. Our cues were peripheral and near the target location, so they could have an exogenous effect. But they were also 100% valid (predictive of the target location), so they could have an endogenous effect. Roach & Hogben (2008) manipulated several factors in this paradigm to probe the role of endogenous versus exogenous attention. First, they manipulated the validity of the cue (how often it appears near the target). When the cue's location was totally random, there was no cueing effect, contrary to the prediction of an exogenous mechanism. They also varied the timing of the cue relative to the stimulus. The cueing effect was not transient, but rather was sustained and increased with more time between the cue and stimulus. The primary mechanism of the attentional effect in this paradigm is therefore endogenous.

Moreover, because in our experiments the cue appeared only 33 ms before the target array, there was not enough time for either endogenous or exogenous attention to focus on the cued location before the stimuli appear. Rather, endogenous attentional selection was likely based on

stimulus representations held in a short-term memory trace (Roach & Hogben, 2008). Therefore, for a participant to benefit from the cue, they must be able to detect and localize the cue, understand the information it conveys, and efficiently filter out information that was seen at all other locations before making a decision. Given that we made the cue especially salient, it is unlikely that detection and localization are what differ between good and poor readers and between children and adults. Rather, our results reveal differences in the capacity to capitalize on the cue's information to select the relevant stimulus.

#### 4.5 Conclusion

In conclusion, children and adolescents with dyslexia are less able than their peers to select information from a task-relevant visual field location and filter out irrelevant information. That attentional skill improves along with fine discrimination ability up until about age 20. Individual differences in covert spatial attention are important not just in reading, but in any daily activity that requires finding objects that are not physically salient or ignoring irrelevant objects that are salient. The biological and environmental causes of the attentional deficits in dyslexia, as well as the cascading effects they may have on other cognitive functions throughout development, are all worthy of further investigation.

#### **Author contributions**

Alex L. White: conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing – original draft, visualization.

Geoffrey M. Boynton: conceptualization, writing – review & editing, supervision, resources.

Jason D. Yeatman: conceptualization, methodology, resources, writing – review & editing, supervision, project administration, funding acquisition.

#### **Data and code availability**

The experiment code, data, and MATLAB code to reproduce each figure and statistic, are all publicly available at the following website: [https://github.com/alexlwhite/WhiteBoyntonYeatman2019\\_Repository](https://github.com/alexlwhite/WhiteBoyntonYeatman2019_Repository) (DOI: 10.5281/zenodo.3361729).

### **Acknowledgements**

This work was funded by: National Eye Institute grants EY026785 and EY029366 to A.L.W.; Microsoft, the Jacobs Foundation Early Career Fellowship, NSF BCS #1551330, and NICHD grant R21HD092771 to J.D.Y; and NEI grant EY12925 to G.M.B. We are grateful to Patrick M. Donnelly for invaluable assistance in data collection, and to John Palmer and Michael Grubb for insightful comments on the manuscript.

### **Conflict of Interest statement**

The authors have no conflicts of interest to declare.

## References

- Akhtar, N., & Enns, J. T. (1989). Relations between covert orienting and filtering in the development of visual attention. *Journal of Experimental Child Psychology, 48*, 315–334.
- Baldassi, S., & Burr, D. C. (2000). Feature-based integration of orientation signals in visual search. *Vision Research, 40*, 1293–300.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological), 90*, 289–300.
- Boden, C., & Giaschi, D. (2007). M-stream deficits and reading-related visual processes in developmental dyslexia. *Psychological Bulletin, 133*, 346–366.
- Bosse, M.-L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition, 104*, 198–230.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 443–446.
- Brodeur, D. A., & Boden, C. (2000). The effects of spatial uncertainty and cue predictability on visual orienting in children. *Cognitive Development, 15*, 367–382.
- Brodeur, D. A., & Enns, J. T. (1997). Covert visual orienting across the lifespan. *Canadian Journal of Experimental Psychology, 51*, 20–35.
- Callens, M., Whitney, C., Tops, W., & Brysbaert, M. (2013). No deficiency in left-to-right processing of words in dyslexia but evidence for enhanced visual crowding. *Quarterly Journal of Experimental Psychology, 66*, 1803–1817.
- Cameron, E. L., Tai, J. C., & Carrasco, M. (2002). Covert attention affects the psychometric function of contrast sensitivity. *Vision Research, 42*, 949–67.
- Cameron, E. L., Tai, J. C., Eckstein, M. P., & Carrasco, M. (2004). Signal detection theory applied to three visual search tasks--identification, yes/no detection and localization. *Spatial Vision, 17*, 295–325.
- Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research, 51*, 1484–1525.
- Carrasco, M., & Yeshurun, Y. (1998). The contribution of covert attention to the set-size and eccentricity effects in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 673–92.
- Cassim, R., Talcott, J. B., & Moores, E. (2014). Adults with dyslexia demonstrate large effects of crowding and detrimental effects of distractors in a visual tilt discrimination task. *PloS One, 9*, e106191.
- Demb, J. B., Boynton, G. M., Best, M., & Heeger, D. J. (1998). Psychophysical evidence

- for a magnocellular pathway deficit in dyslexia. *Vision Research*, 38, 1555–1559.
- Demb, J. B., Boynton, G. M., & Heeger, D. J. (1997). Brain activity in visual cortex predicts individual differences in reading performance. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 13363–13366.
- Eden, G. F., Stein, J. F., & Wood, F. B. (1994). Differences in eye movements and reading problems in dyslexic and normal children. *Vision Research*, 34, 1345–1358.
- Ekstrand, C., Neudorf, J., Gould, L., Mickleborough, M., & Borowsky, R. (2019). Where words and space collide: The overlapping neural activation of lexical and sublexical reading with voluntary and reflexive spatial attention. *Brain Research*, 1706, 1–12.
- Enns, J. T., & Brodeur, D. A. (1989). A developmental study of covert orienting to peripheral visual cues. *Journal of Experimental Child Psychology*, 48, 171–189.
- Facoetti, A., Paganoni, P., Turatto, M., Marzola, V., & Mascetti, G. G. (2000). Visual-spatial attention in developmental dyslexia. *Cortex*, 36, 109–123.
- Folk, C. L., & Hoyer, W. J. (1992). Aging and shifts of visual spatial attention. *Psychology and Aging*, 7, 453–465.
- Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K., & Facoetti, A. (2012). A Causal Link between Visual Spatial Attention and Reading Acquisition. *Current Biology*, 22, 814–819.
- Gaspelin, N., Margett-Jordan, T., & Ruthruff, E. (2015). Susceptible to distraction: Children lack top-down control over spatial attention capture. *Psychonomic Bulletin and Review*, 22, 461–468.
- Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The Neural Basis of Perceptual Learning. *Neuron*, 31, 681–697.
- Goldberg, M. C., Maurer, D., & Lewis, T. L. (2001). Developmental changes in attention: the effects of endogenous cueing and of distractors. *Developmental Science*, 4, 209–219.
- Gori, S., & Facoetti, A. (2015). How the visual aspects can be crucial in reading acquisition: The intriguing case of crowding and developmental dyslexia. *Journal of Vision*, 15, 1–20.
- Gori, Simone, Seitz, A. R., Ronconi, L., Franceschini, S., & Facoetti, A. (2015). Multiple Causal Links Between Magnocellular–Dorsal Pathway Deficit and Developmental Dyslexia. *Cerebral Cortex*, bhv206.
- Hartley, A. A., Kieley, J. M., & Slabach, E. H. (1990). Age Differences and Similarities in the Effects of Cues and Prompts. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 523–537.



- Hawelka, S., & Wimmer, H. (2005). Impaired visual processing of multi-element arrays is associated with increased number of eye movements in dyslexic reading. *Vision Research*, *45*, 855–863.
- Iarocci, G., Enns, J. T., Randolph, B., & Burack, J. A. (2009). The modulation of visual orienting reflexes across the lifespan. *Developmental Science*, *12*, 715–724.
- Joo, S. J., White, A. L., Strodman, D. J., & Yeatman, J. D. (2018). Optimizing text for an individual's visual system: The contribution of crowding to reading difficulties. *Cortex*, *103*, 291–301.
- Kevan, A., & Pammer, K. (2008). Making the link between dorsal stream sensitivity and reading. *NeuroReport*, *19*, 467–470.
- Landry, O., Johnson, K. A., Fleming, S. J., Crewther, S. G., & Chouinard, P. A. (2019). A new look at the developmental profile of visual endogenous orienting. *Journal of Experimental Child Psychology*, *183*, 158–171.
- Leclercq, V., & Siéroff, E. (2013). Development of endogenous orienting of attention in school-age children. *Child Neuropsychology*, *19*, 400–419.
- Legge, G. E., Cheung, S.-H. S. H., Yu, D., Chung, S. T. L., Lee, H.-W. H. W., & Owens, D. P. (2007). The case for the visual span as a sensory bottleneck in reading. *Journal of Vision*, *7*, 1–15.
- Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition*, *6*, 312–319.
- Lu, Z. L., & Doshier, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research*, *38*, 1183–1198.
- Maunsell, J. H. R. (2015). Neuronal Mechanisms of Visual Attention. *Annual Review of Vision Science*, *1*, 373–391.
- McDougall, P., Borowsky, R., MacKinnon, G. E., & Hymel, S. (2005). Process dissociation of sight vocabulary and phonetic decoding in reading: A new perspective on surface and phonological dyslexias. *Brain and Language*, *92*, 185–203.
- Michael, G. A., Lété, B., & Ducrot, S. (2013). Trajectories of attentional development: An exploration with the master activation map model. *Developmental Psychology*, *49*, 615–631.
- Moores, E., Cassim, R., & Talcott, J. B. (2011). Adults with dyslexia exhibit large effects of crowding, increased dependence on cues, and detrimental effects of distractors in visual search tasks. *Neuropsychologia*, *49*, 3881–3890.
- Morgan, M. J., Ward, R. M., & Castet, E. (1998). Visual Search for a Tilted Target: Tests of Spatial Uncertainty Models. *Quarterly Journal of Experimental Psychology*, *51A*, 347–370.

- Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research*, 34, 1703–1721.
- Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 108–130.
- Pammer, K., Hansen, P., Holliday, I., & Cornelissen, P. (2006). Attentional shifting and the role of the dorsal pathway in visual word recognition. *Neuropsychologia*, 44, 2926–2936.
- Pearson, D. A., & Lane, D. M. (1990). Visual Attention Movements: A Developmental Study. *Child Development*, 61, 1779–1795.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Peterson, R. L., & Pennington, B. F. (2015). Developmental Dyslexia. *Annual Review of Clinical Psychology*, 11, 283–307.
- Petrov, A. a, Doshier, B. A., & Lu, Z.-L. (2005). The dynamics of perceptual learning: an incremental reweighting model. *Psychological Review*, 112, 715–43.
- Plude, D. J., Enns, J. T., & Brodeur, D. A. (1994). The development of selective attention: A life-span overview. *Acta Psychologica*, 86, 227–272.
- Posner, M. I. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology*, 32, 3–25.
- Prins, N., & Kingdom, F. A. A. (2009). Palamedes: Matlab routines for analyzing psychophysical data.
- Rayner, K. (1985). Do faulty eye movements cause dyslexia? *Developmental Neuropsychology*, 1, 3–15.
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62, 1457–1506.
- Reynolds, J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annual Review of Neuroscience*, 27, 611–647.
- Ristic, J., & Enns, J. T. (2015). Attentional Development. *Handbook of Child Psychology and Developmental Science*, 1–45.
- Ristic, J., & Kingstone, A. (2009). Rethinking attentional development: Reflexive and volitional orienting in children and adults. *Developmental Science*, 12, 289–296.
- Roach, N. W., & Hogben, J. H. (2004). Attentional modulation of visual processing in adult dyslexia. *Psychological Science*, 15, 650–654.

- Roach, N. W., & Hogben, J. H. (2007). Impaired filtering of behaviourally irrelevant visual information in dyslexia. *Brain, 130*, 771–785.
- Roach, N. W., & Hogben, J. H. (2008). Spatial cueing deficits in dyslexia reflect generalised difficulties with attentional selection. *Vision Research, 48*, 193–207.
- Schul, R., Townsend, J., & Stiles, J. (2003). The development of attentional orienting during the school-age years. *Developmental Science, 6*, 262–272.
- Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *The New England Journal of Medicine, 326*, 145–150.
- Skottun, B. C. (2000). The magnocellular deficit theory of dyslexia: The evidence from contrast sensitivity. *Vision Research, 40*, 111–127.
- Torgesen, J., Rashotte, C., & Wagner, R. (1999). TOWRE-2: Test of Word Reading Efficiency, 2nd Ed. *Austin, TX: Pro-Ed*.
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45*, 2–40.
- Vidyasagar, T. R., & Pammer, K. (1999). Impaired visual search in dyslexia relates to the magnocellular pathway in attention. *Neuroreport, 10*, 1283–1287.
- Vidyasagar, T. R., & Pammer, K. (2010). Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. *Trends in Cognitive Sciences, 14*, 57–63.
- Wainwright, A., & Bryson, S. E. (2002). The development of exogenous orienting: Mechanisms of control. *Journal of Experimental Child Psychology, 82*, 141–155.
- Wainwright, A., & Bryson, S. E. (2005). The Development of Endogenous Orienting : Control Over the Scope of Attention and Lateral Asymmetries. *Developmental Neuropsychology, 27*, 237–255.
- White, A. L., Palmer, J., & Boynton, G. M. (2018). Evidence of serial processing in visual word recognition. *Psychological Science, 29*, 1062–1071.

## Supplemental Material for:

The link between reading ability and visual spatial attention across development

by White, Boynton & Yeatman

### Data and code availability

The data collected for this study, alongside code to run the experiment and reproduce each figure and statistic, are all available at the following URL: [https://github.com/alexwhite/WhiteBoyntonYeatman2019\\_Repository](https://github.com/alexwhite/WhiteBoyntonYeatman2019_Repository) (DOI: [10.5281/zenodo.3361729](https://doi.org/10.5281/zenodo.3361729)). The raw trial-level data are in the form of text files, and the experiment and analysis code are written in Matlab.

### Experiment versions

Over the course of data collection, we used three slightly different versions of the experiment that differ primarily in which conditions were included (**Table S1**). All versions included the Uncued (U) and Cued (C) conditions. Version 1 also included the Single Stimulus (S) condition, and we ran 2 blocks of each of the three conditions in each session. Version 2 included only the most critical conditions: Uncued and Cued (3 blocks each). Version 3 additionally included the Single Stimulus and Small Cue (Cs) conditions (2 blocks each). Children under 14 (except for one 8-year old, by mistake) were assigned to Versions 1 or 2. Version 2 was created for the second half the sample of children, in order to focus testing in a short amount of time on the two most critical conditions. Version 3 was created to include the Small Cue and more closely match the design of Roach & Hogben (2007) in adults. All participants aged 14 or older participated in Version 3.

Version	Block order	# Participants	Age range
1	[S U C U C S]	26	6-13
2	[U C U C U C]	19	5-13
3	[S U C Cs C Cs U S] or [Cs C U S U S C Cs]	78	14-70*

**Table S1:** The three experiment versions. Condition labels: U = Uncued; C = Cued; Cs = Small Cue; S = Single Stimulus. “Easy” trials were the minority of trials in each block when the staircase was not updated, and the tilt level set much higher than the staircase’s current threshold estimate in order to improve our fits of the upper asymptote of the psychometric function. \*One 8-year old participant did Version 3, but all others were at least 14 years old.

## Theoretical motivation for the cueing effect as a difference in log thresholds

We built a computational model to verify that the difference in log thresholds is an appropriate measure of the effect of spatial attention elicited by the cue. The model assumes that the visual system estimates the orientation  $\theta$  of each Gabor stimulus in the display, where 0 is vertical. Each  $\theta$  is a random draw from a normal distribution with standard deviation  $\sigma$  and mean centered on the true orientation. The decision rule is to report “clockwise” if the target  $\theta > 0$ , and “counterclockwise” otherwise. Single stimulus thresholds are controlled by  $\sigma$ . In the Uncued and Cued conditions (both with set size 8), a capacity or resource limitation can be simulated by increasing  $\sigma$  by a factor  $\kappa$ .

The model identifies the target as item with the largest absolute value of  $\theta$  (greatest deviation from vertical). In the Uncued condition, all eight items are given equal weight, and due to noise in the perceptual process, distractors can be mistaken for the target. In the Cued condition, the influence of distractors is reduced by multiplying their  $\theta$  values by an ‘exclusion factor’  $\xi < 1$ . The benefit of the cue is maximized if  $\xi = 0$ , in which case only the cued item influences the decision.

We obtained thresholds from the model by simulating thousands of trials and varying the target tilt ( $\mu$ ), computing proportion correct, and fitting psychometric functions.

The difference of log thresholds (Uncued-Cued) varied linearly with the exclusion factor  $\xi$ . That relationship was robust to variations in  $\sigma$  and  $\kappa$ . Therefore, using this metric of the cueing effect, we can assess individual differences in the efficiency of filtering out noise from uncued distractors.

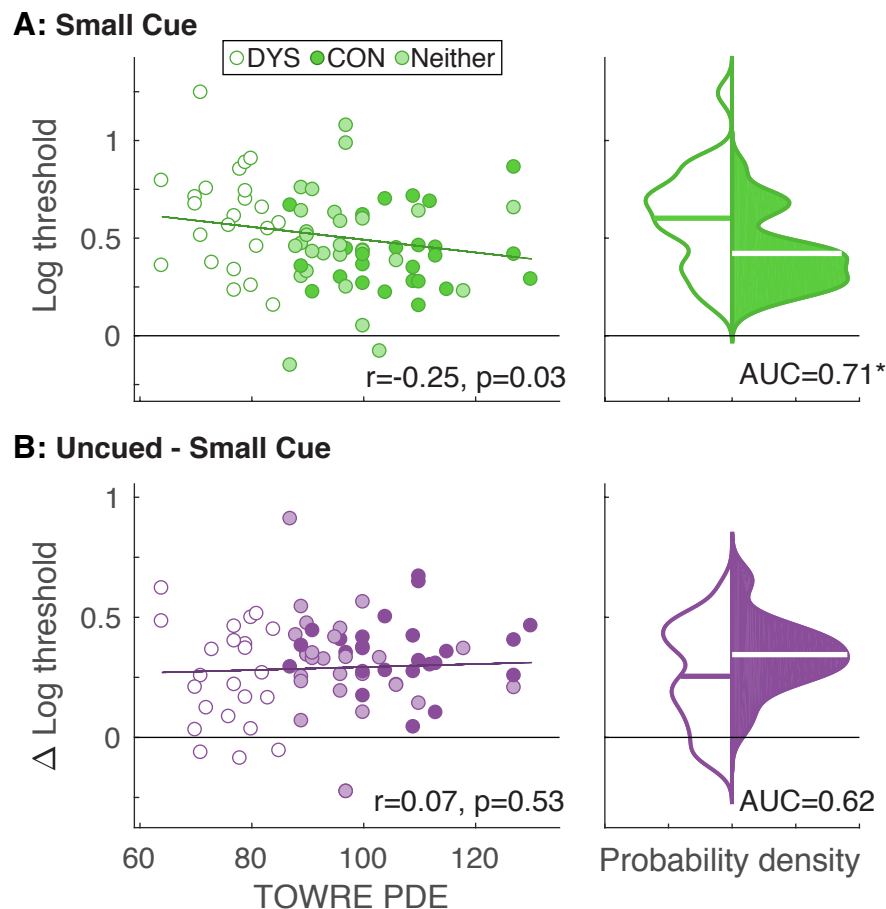
We obtained similar results when using a pooling rule rather than a max rule: the target is not specifically localized, but tilt is estimated by taking a weighted average of all 8 items. To simulate the cueing benefit, the parameter  $\xi$  lowers the weight of distractors in the computation of the average tilt. Again, the difference of log thresholds is linearly related to  $\xi$ .

This model did not address the possibility of *signal enhancement* as an attentional mechanism distinct from noise exclusion. Signal enhancement could be accomplished by reducing  $\sigma$  for the cued item, in addition to excluding distractor noise. We do not rule out signal enhancement, although Roach & Hogben (2007, 2008) concluded that distractor exclusion is the primary mechanism in this task. Moreover, we have no reason to expect that some signal enhancement would not also be captured by the difference of log thresholds.

## Results from the Small Cue condition

77 participants over the age of 14 were tested in Version 3 of the experiment, which included the Small Cue condition. One was excluded because their accuracy was not above chance in this condition. 23 of the included participants belonged to the DYS group and 25 to the CON group.

Thresholds in the two cued conditions (large and small) were highly correlated ( $r=0.68$ ,  $p<10^{-11}$ ), and on average were slightly higher in the Small Cue condition ( $3.7^\circ$  for the small cue vs.  $3.3^\circ$  for the large cue;  $t(75) = 3.24$ ,  $p=0.002$ ). Therefore, although the Small Cue was less salient, it seems that the two cues had very similar effects within individuals. They may nonetheless interact with reading ability in different ways.



**Figure S1.** Association between performance in the small cue condition and reading ability. Panel A is structured in the same way as Figure 3, and Panel B as Figure 4. One difference that here we plot raw thresholds in (A) and cueing effects on thresholds in (B), whereas the previous plot *residuals* from the developmental model fit. In this restricted sample we there was no significant effect of age in a developmental model.

Because this sample was smaller and the age range constrained, we could not find a clear developmental pattern in Small Cue thresholds. The initial slope of a piecewise linear model

was not significantly different from 0, and a simpler linear model of age found no significant effect. Similarly, there was no clear effect of age on the cueing effect (threshold difference Uncued – Small Cued). In the analysis presented below, therefore, we do not divide into age bins nor analyze the distributions of residuals from the age model. But we do include age as a covariate in the statistical models. The left panel of **Figure S1A** plots the individual thresholds in the Small Cue condition as a function of reading score (TOWRE PDE). There was a significantly negative correlation ( $r=-0.25$ ,  $p=0.03$ ). We further analyzed these thresholds with a linear model that included age, reading score, ADHD diagnosis, and non-verbal IQ. No single parameter was significant, including reading score ( $t(67) = 1.66$ ,  $p=0.10$ ). However, the beta weight describing the relationship between reading score and threshold ( $-0.003$ ) was similar to that for the larger salient cue ( $-0.004$ ). The solid line in **Figure S1A (left)** is the prediction based on reading score alone.

We used a similar model with the same covariates to assess threshold differences between the DYS and CON groups (rather than reading score). Thresholds in the DYS group (mean =  $0.60 \pm 0.05 \log_{10}(\text{deg})$ ) were higher than in the CON group ( $0.42 \pm 0.04$ ;  $t(40) = 2.33$ ,  $p=0.025$ ). The discriminability of the groups was assessed with an ROC analysis (right panel of **Fig. S1A**). The AUC was 0.71, significantly greater than 0.5 (null distribution 95%CI = [0.34 0.66]).

Finally, we estimated the *cueing effect*: the difference in thresholds Uncued - Small Cue. Individual cueing effects are plotted in **Figure S1B** as a function of reading score. Those two measures were not correlated ( $r=0.07$ ,  $p=0.53$ ), and the slope of the linear effect (0.0006) was not different from 0 ( $t(67)<1$ ). The DYS group had a smaller effect (mean =  $0.255 \pm 0.04$ ) than the CON group ( $0.345 \pm 0.03$ ). But according to a linear model with covariates for normalized age, non-verbal IQ, and ADHD diagnosis, the two groups did not quite differ significantly ( $t(40) = 1.75$ ,  $p=0.088$ ). None of the covariates had significant effects either. The ROC analysis yielded a similarly small effect: AUC = 0.62, null 95% CI = [0.33 0.66].

When analyzing only participants 20 years of age or older, accuracy was higher but still not above chance: AUC = 0.67 (null 95% CI = [0.30 0.70]). In contrast, ROC analysis of the cueing effect from the bigger red cue (over age 20) yielded an AUC of 0.60 (null 95%CI = [0.30 0.70]).

## Lapse rates

The  $\lambda$  parameter in our psychometric function fits is the difference of the upper asymptote from 1, and is related to lapse rate. Specifically,  $p(\text{lapse}) = 2\lambda$ .  $\lambda$  tended to decrease with age, although several of the youngest participants had  $\lambda=0$ .

The average  $\lambda$  in participants under 11 years of age was  $0.086 \pm 0.007$ ; among participants over 20 it was  $0.043 \pm 0.004$ . Fitting the same piecewise linear model as we did for thresholds, we found a decline in  $\lambda$  from ages 5 to 19 (slope =  $-0.003/\text{year}$ ), and then a plateau at  $\lambda = 0.04$ .

We then analyzed the residuals of this developmental model of lapse rates, to see how they relate to reading ability. The TOWRE PDE score did not predict  $\lambda$ , when controlling for age, non-verbal IQ and ADHD diagnosis (estimated slope = 0.001,  $t(112) = 0.47$ ,  $p=0.64$ ). Moreover, the  $\lambda$  residuals did not differ between the DYS group (mean =  $0.002 \pm 0.005$ ) and the CON group ( $-0.003 \pm 0.006$ ); AUC = 0.54 (null CI = [0.37 0.63]).

## Response times

For each participant, in each condition, we computed the geometric mean of response times (RTs) in each condition. We then analyzed them in a similar way as we analyzed thresholds.

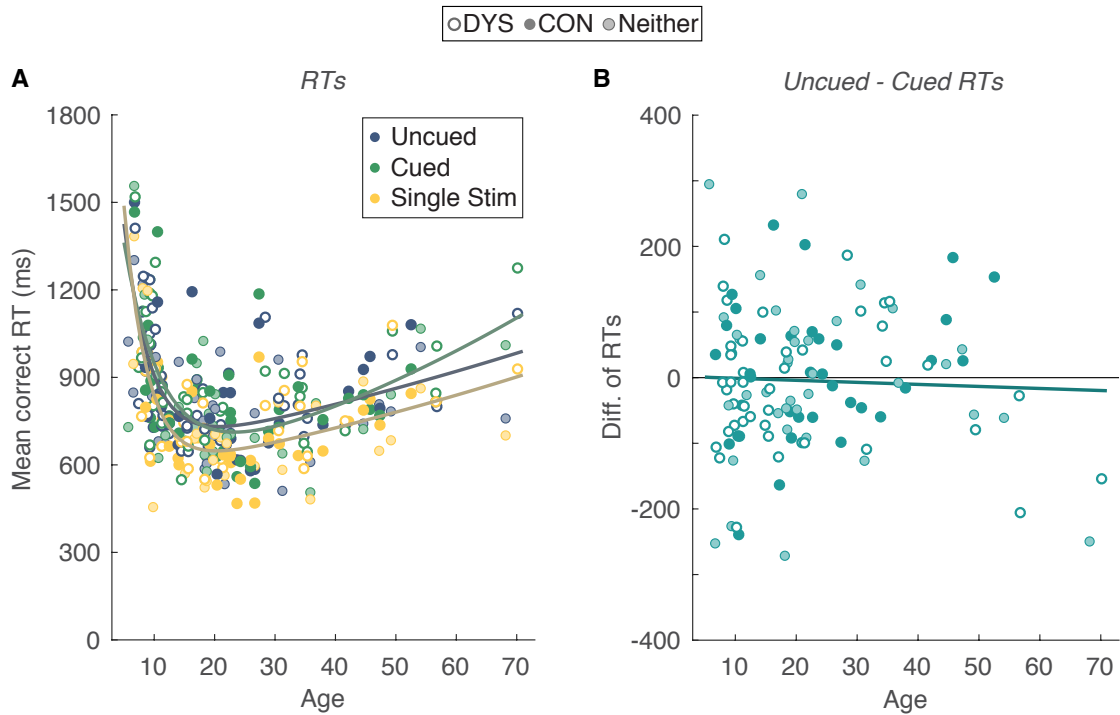
The developmental time-course on RTs was best fit with an asymmetric u-shaped function of the form:

$$RT = s(a_1e^{-tb_1} + a_2e^{tb_2})$$

where  $RT$  is response time in ms and  $t$  is age in years. The parameter  $s$  is a scale factor.  $a_1$  and  $b_1$ , and  $a_2$  and  $b_2$ , together control the steepness of the downwards and upwards portion of the curve, respectively.

The best fits to each stimulus condition are shown in **Figure S2A**. None of the parameters differed significantly between conditions. The cueing effect on RTs (Uncued – Cued) for individual observers are shown in **Figure S2B**, along with a piecewise linear model in which the second line was allowed to have a slope different from 0. There was on average very little cueing effect, and no significant change over age.

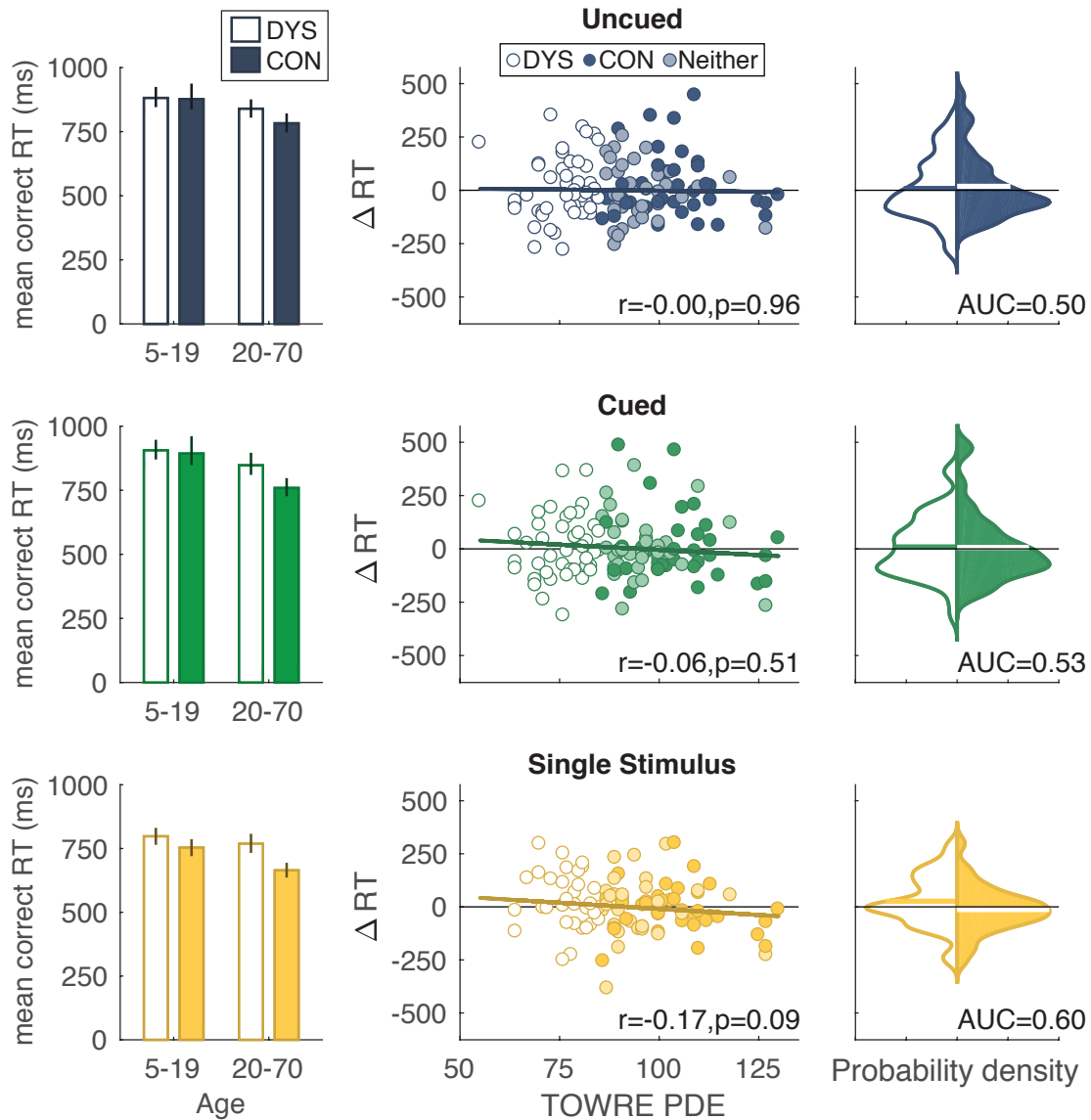




**Figure S2:** Developmental patterns on correct response times (RTs) and cueing effects on RTs. **(A)** RTs as a function of age in each of three conditions. Each point is one participant’s mean correct RT in one condition. The fill color of each point indicates the participant’s reading ability group, as indicated by the legend at top. The lines are the best-fitting u-shaped functions. **(B)** The difference in mean RTs between the Uncued and Cued conditions as a function of age. Each dot is one participant’s difference. The line is the best-fitting piecewise linear function.

In **Figure S3**, we divide up the participants by reading ability. The left column shows mean RTs in two age bins (below and above age 20) further sorted in the DYS and CON groups. The DYS group tended to have slower RTs, especially in the older age bin. That pattern raises an interesting hypothesis: older participants with dyslexia have thresholds similar to their age-matched peers (unlike the younger DYS group) because they have learned to respond more slowly and carefully. However, there were no significant effects of reading ability or interaction with age group or condition.

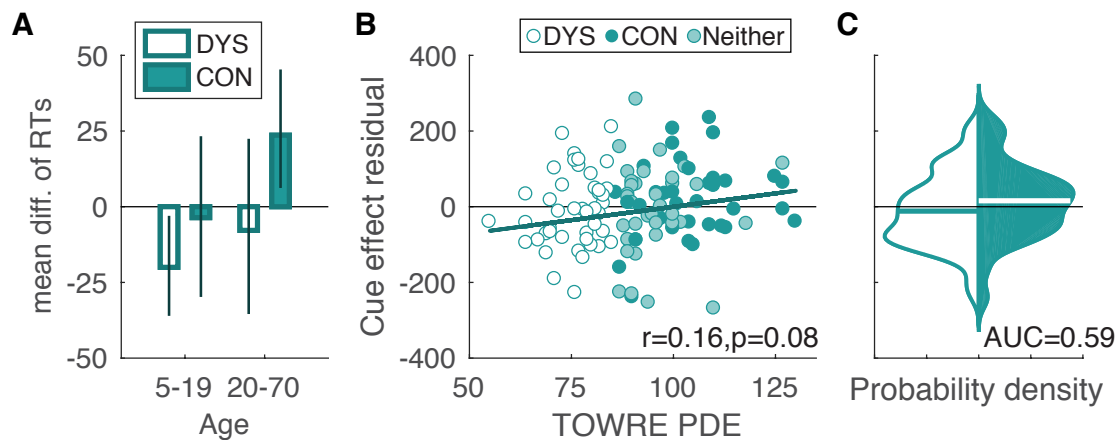
The middle column of Figure S3 plots the association between reading ability (TOWRE PDE score) and the RT *residuals* from the developmental models in Figure SA. In no condition was there a significant correlation, nor did a linear model find a slope significantly different from 0 (controlling for ADHD and non-verbal IQ). The ROC analysis (Fig. S3, right column) similarly found no difference between DYS and CON groups in terms of RT residuals.



**Figure S3.** The effect of reading ability on response times. **Left column:** mean correct RTs for subjects divided into two age groups (left pair of bars: 5-19 years; right pair of bars: 20-70 years) and two reading groups (DYS and CON). Error bars are bootstrapped 68% CIs. **Middle column:** correlations between TOWRE phonemic decoding score (PDE) and the residuals of the developmental model plotted in Figure S2. This includes *all* subjects. The y-axis is the difference between each subject's mean RT and the RT predicted by the developmental model. The solid line is the prediction based on PDE score from a model that also included ADHD and non-verbal IQ as predictors. **Right column:** smoothed probability density functions of the residuals plotted in the middle column, just for the DYS and CON groups. The horizontal line superimposed on each distribution is the mean. The AUC is the area under the ROC curve. No AUC differed significantly from chance (0.5).

We would have reason for concern if poor readers responded more quickly than good readers, which could explain why they have higher thresholds; that would suggest a difference in decision-making strategy rather than perceptual sensitivity. That did not occur, however; the trend was for slower responses in poor readers, but overall RTs in this task are not strongly associated with reading ability.

Finally, in **Figure S4** we evaluate the association between reading ability and the cueing effect on RTs. In **Figure S4A** we see that both for participants under and over age 20, the DYS group tended to have a larger (more positive) cueing effect than the CON group by about 24 ms (when controlling for ADHD and non-verbal IQ). However, that difference was not significant ( $F(1,69)=0.93$ ,  $p=0.34$ ), and it did not interact with age group ( $F(1,69)<1$ ).



**Figure S4:** Relation between the cueing effect on RTs and reading ability. **A:** mean cueng effects, expressed as the difference in RT between the Uncued and Cued conditions. Formatted as in Fig. S3 (left column). **B:** Correlation between reading score and the residuals from the developmental model of the cueing effect on RTs. The solid line is the prediction based on PDE score from a model that also included ADHD and non-verbal IQ as predictors. **C:** Smoothed probability density functions of the residuals plotted in the middle column, just for the DYS and CON groups. The horizontal line superimposed on each distribution is the mean.

When including all participants, there was small positive correlation between reading score and the residuals of the cueing effect ( $r=0.16$ ,  $p=0.08$ ). The linear model slope (controlling for ADHD and non-verbal IQ) was positive (1.41;  $t(108) = 2.06$ ,  $p=0.042$ ). The AUC for discriminating the DYS from the CON group on the basis of their RT cueing effects was 0.59 (not significantly above chance; null 95% CI = [0.38 0.62]).

In conclusion, there was overall very little effect of the cue on RTs, but good readers tended to have larger cueing effects. That is consistent with the patterns on thresholds.

### Analysis using sight word efficiency as the reading ability measure

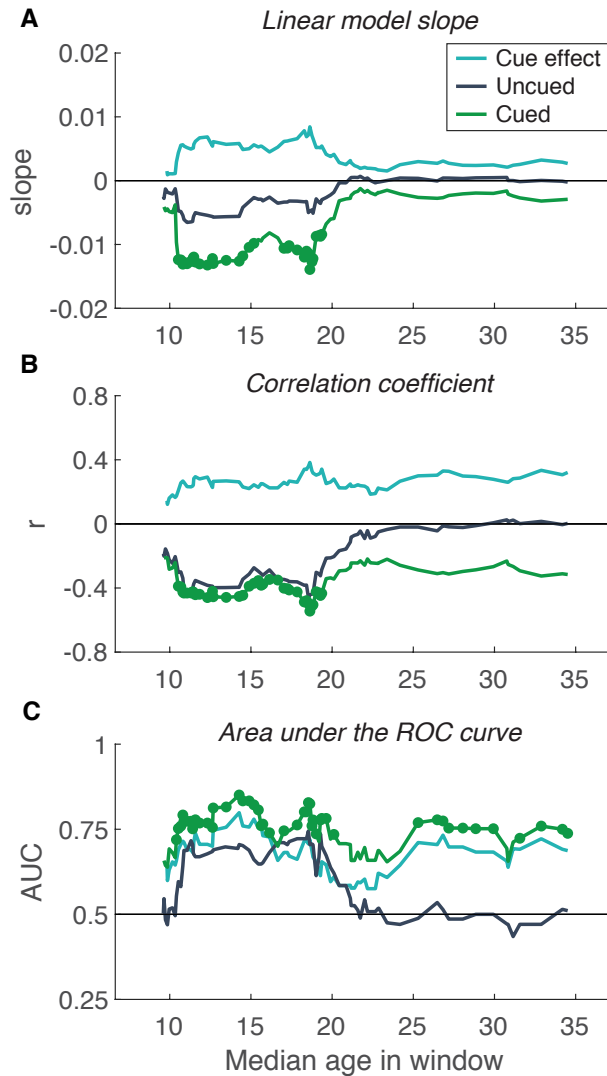
In the main text, we used the TOWRE phonemic decoding efficiency (PDE) score as the measure of reading ability. We also conducted the same analyses using the TOWRE sight word efficiency (SWE) score instead (Torgesen, Rashotte, & Wagner, 1999).

The two measures (PDE and SWE) were highly correlated ( $r=0.75$ ). However, our participants were sorted differently into the DYS and CON groups, depending on which reading score we use. 28 fell into both DYS groups. Only 12 met the criteria based on SWE but not on PDE, and 18 based on PDE but not on SWE. We therefore analyzed the set of 40 participants that met the criteria based on SWE (score  $\leq 85$  and history of reading problems), regardless of their PDE score.

Overall, the results were quite similar when using the SWE rather than the PDE. Across the whole sample, the correlations with cued thresholds were identical ( $r=-0.24$ ), as was the area under the ROC curve for discriminating the DYS and CON groups (AUC=0.66). The correlations with the cueing *effect* were similar ( $r=0.17$  for SWE,  $r=0.15$  for PDE), but the AUC was slightly smaller for SWE (0.6 vs. 0.64).

The one interesting difference is that for SWE, the associations with cued thresholds and cueing effects do not seem to decrease with age as much as they do for PDE. That is summarized with the sliding window analysis in **Figure S5** (analogous to Figure 5 in the main text). The pattern of linear models slopes relating SWE to thresholds (Fig. S5A) is quite similar to the pattern for PDE (Fig. 5A). The correlation coefficients between SWE and cued thresholds and cueing effects (Fig. S5B) do not fall as close to zero after age 20 as they do with PDE. That can also be seen in the AUC values for discriminating DYS from CON groups (Fig. S5C). The AUCs for cued thresholds remain statistically significant well after age 20, in contrast to what was observed when sorting the groups based on PDE score (Fig. 5C).

That pattern raises an interesting hypothesis for future research: among adults with dyslexia, the relative deficiency in covert spatial attention is stronger for the subset of individuals who are impaired in real word reading, not (just) phonemic decoding.



**Figure S5.** Association between reading ability, assessed with the TOWRE SWE score, and task performance assessed in sliding windows across the age range. **(A)** Slopes of a linear fit between SWE scores and *residuals* of the developmental model for Uncued thresholds, Cued thresholds, and the Uncued - Cued effect. The linear model included ADHD diagnosis and non-verbal IQ as covariates. Each time window included 1/3 of the sample (N=40). **(B)** Similar analysis for the linear correlation coefficients between residuals and TOWRE SWE. **(C)** Similar analysis of the Area under the ROC curve that distinguishes the DYS group from the CON group. In all three panels, dots on individual time-points indicate that the slope significantly different from 0 (corrected for multiple comparisons).

### Supplemental References

Roach, N. W., & Hogben, J. H. (2007). Impaired filtering of behaviourally irrelevant visual information in dyslexia. *Brain, 130*, 771–785.

Roach, N. W., & Hogben, J. H. (2008). Spatial cueing deficits in dyslexia reflect generalised difficulties with attentional selection. *Vision Research, 48*, 193–207.

Torgesen, J., Rashotte, C., & Wagner, R. (1999). TOWRE-2: Test of Word Reading Efficiency, 2nd Ed. *Austin, TX: Pro-Ed.*