# Excessive visual crowding effects in developmental dyscalculia

**Elisa Castaldi***

Department of Translational Research and New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy
Department of Neuroscience, Psychology, Pharmacology and Child Health, University of Florence, Florence, Italy
Cognitive Neuroimaging Unit, NeuroSpin Center, CEA DRF/JOLIOT, INSERM, Université Paris-Saclay, Gif-sur-Yvette, Paris, France ✉

**Marco Turi***

Fondazione Stella Maris Mediterraneo, Potenza, Italy ✉

**Sahawanatou Gassama**

Paris Santé Réussite, Centre de diagnostic des troubles des apprentissages, Paris, France ✉

**Manuela Piazza**

Center for Mind/Brain Sciences, University of Trento, Italy ✉

**Evelyn Eger**

Cognitive Neuroimaging Unit, NeuroSpin Center, CEA DRF/JOLIOT, INSERM, Université Paris-Saclay, Gif-sur-Yvette, Paris, France ✉

**Visual crowding refers to the inability to identify objects when surrounded by other similar items. Crowding-like mechanisms are thought to play a key role in numerical perception by determining the sensory mechanisms through which ensembles are perceived. Enhanced visual crowding might hence prevent the normal development of a system involved in segregating and perceiving discrete numbers of items and ultimately the acquisition of more abstract numerical skills. Here, we investigated whether excessive crowding occurs in developmental dyscalculia (DD), a neurodevelopmental disorder characterized by difficulty in learning the most basic numerical and arithmetical concepts, and whether it is found independently of associated major reading and attentional difficulties. We measured spatial crowding in two groups of adult individuals with DD and control subjects. In separate experiments, participants were asked to discriminate the orientation of a Gabor patch either in isolation or under spatial crowding. Orientation discrimination thresholds were comparable across groups when stimuli were shown in isolation, yet they were much higher for the DD group with respect to the control group when the target was crowded by closely neighbouring flanking gratings. The difficulty in discriminating orientation (as reflected by the combination of accuracy and reaction times) in the DD compared to the control group persisted over several larger target flanker distances. Finally, we found that the degree of such spatial crowding correlated with impairments in mathematical abilities even when controlling for visual attention and reading skills. These results suggest that excessive crowding effects might be a characteristic of DD, independent of other associated neurodevelopmental disorders.**

## Introduction

Developmental dyscalculia (DD) is a neurodevelopmental learning disability characterized by difficulty in learning about numbers and arithmetic, which manifests in children despite adequate neurological development, intellectual abilities and schooling opportunity (American Psychiatric Association, 2013). DD affects a wide range of mathematical abilities: DD individuals can have difficulties in understanding the meaning of numerical magnitudes and Arabic digits, in retrieving arithmetic facts from memory and in automatizing simple calculation procedures. Due to the variety of processes found to be often impaired, DD has been conceptualized as a multidimensional disorder: the difficulties in mathematical competence may

emerge from the combination of weak domain-general functions (such as attention, memory and cognitive control) and domain-specific impairments in mastering numerical and arithmetical concepts (Fias, Menon, & Szucs, 2013; Fias, 2016; Iuculano, 2016).

Visuospatial working memory deficits occur in pure DD and these are stronger compared to profiles with associated dyslexia which are instead characterized by stronger verbal working memory deficits (Szűcs, 2016). Impairments in visuo-spatial attention and alertness have been reported in DD individuals, despite not meeting the criteria for attention-deficit/hyperactivity disorder (ADHD) (Askenazi & Henik, 2010). During numerical processing, calculation and math problem solving, these domain-general functions are thought to interact with a specific 'core knowledge' of magnitude, foundational for mathematical competence. Since very early in life, humans can perceive changes in the number of objects in an image (Brannon, Abbott, & Lutz, 2004; Izard, Sann, Spelke, Streri, & Gallistel, 2009; Libertus, Starr, & Brannon, 2014; de Hevia, 2016; de Hevia, Castaldi, Streri, Eger, & Izard, 2017), an ability that is thought to be based on our 'number sense' (Dehaene, 1997). During development, numerical magnitudes, initially experienced in their non-symbolic format, are thought to be mapped onto their symbolic counterpart, setting the base for formal arithmetical learning (Piazza, 2010). In line with this view, the ability to discriminate between non-symbolic numerical quantities (also known as 'numerical acuity') was found to be predictive of arithmetical skills in the neurotypical population (Halberda, Mazzocco, & Feigenson, 2008; Piazza, Pica, Izard, Spelke, & Dehaene, 2013; Chen & Li, 2014; Anobile, Castaldi, Turi, Tinelli, & Burr, 2016a, 2018) and to refine with education (Piazza et al., 2013). Importantly, these basic numerical abilities were found to be disproportionally impaired in DD individuals: they tend to be slower and/or less accurate than their age-matched peers when briefly shown with two sets of dots or with two digits and asked to choose the numerically larger one or when tested with estimation tasks requiring mapping numbers across formats (Butterworth, 2005; Rousselle & Noël, 2007; Iuculano, Tang, Hall, & Butterworth, 2008; Piazza et al., 2010; Butterworth, 2010; Mejias, Grégoire, & Noël, 2012). In line with these observations, some of the most influential models of DD attributed the origin of the numerical and mathematical difficulties to a specific deficit in the core representation of quantity (Dehaene, Piazza, Pinel, & Cohen, 2003; Butterworth, 2005, 2010).

If DD originates from, among other factors, a weak number sense, then it is important to investigate whether the visual mechanisms supporting the extraction of numerosity from an image are impaired. To date, the exact visual mechanisms mediating numerosity perception are still a matter of debate. Some authors suggested that number is sensed directly through

dedicated "number detectors" (Burr & Ross, 2008; Ross, 2010), while others proposed instead that number is perceived through mechanisms related to texture-density processing (Dakin, Tibber, Greenwood, Kingdom, & Morgan, 2011; Morgan, Raphael, Tibber, & Dakin, 2014). Importantly, behavioural studies in individuals without math difficulties have provided evidence that the mechanism through which ensembles are perceived can change depending on how cluttered items are (Anobile, Cicchini, & Burr, 2016b): Anobile, Cicchini, and Burr (2013) measured discrimination thresholds as a function of dot numerosity and found that, for relatively sparse arrays of smaller numerosites, the thresholds followed Weber's law (remained constant across numbers), however, for larger numerosities when sets became highly cluttered, the discrimination thresholds decreased steadily with numerosity, following a square-root law. The fact that discrimination thresholds obeyed different psychophysical laws suggested that visual arrays can be processed by two independent perceptual systems: one related to number perception, recruited when viewing sufficiently sparse arrays, and the other related to texture-density perception, recruited when individual items are too cluttered to be clearly segregated (Anobile, Turi, Cicchini, & Burr, 2015, 2016b; Burr, Anobile, & Arrighi, 2017). Interestingly, only the ability to discriminate the numerosity of sparse, but not of cluttered arrays, was found to predict numerical skills in typically developing children (Anobile et al., 2016a), suggesting that only the functionality of the system mediating number perception, and not of the one mediating texture-density perception, may be relevant for the development of formal mathematical abilities. When studying the variation of discrimination thresholds as a function of dot numerosity, Anobile et al. (2015) observed that the switching point between the two psychophysical regimes depended on the number of dots per visual degree and on the eccentricity at which the arrays were shown, while being unrelated to the individual items' size. These observations suggested that the recruitment of one of the two systems supporting ensemble perception might be regulated by mechanisms related to visual crowding, a process which sets limits to our ability to identify, locate and count objects when they are cluttered together (for reviews see: Levi, 2008; Pelli, 2008; Pelli & Tillman, 2008; Whitney & Levi, 2011). Crowding degrades visual perception, making individual items appear jumbled together with the surrounding objects (Pelli, Palomares, & Majaj, 2004). When multiple nearby items are merged into a single percept, numerosity might undergo underestimation: Valsecchi, Toscani, and Gegenfurtner (2013) showed that the perceived numerosity of arrays presented in the periphery was reduced compared to central viewing, a phenomenon that could be simulated by a texture synthesis algorithm inducing texture

formation of cluttered items in random dot arrays (Balas, 2016). The perceived numerosity was found to decrease for smaller inter-item distance, and this effect could not be accounted by blurring in peripheral vision nor by misperception of stimulus size, leading Valsecchi et al. (2013) to attribute the underestimation of peripherally viewed arrays to visual crowding (however see: Chakravarthi & Bertamini, 2020, for an alternative interpretation of underestimation based on clustering rather than crowding). Interestingly, crowding may also lead to perception of additional elements in the scene: using a change detection task, Greenwood, Bex, and Dakin (2010) showed that crowing induced participants to perceive a target noise or even an inexistent target (blank space) to be oriented as the flanking gratings.

If mastering non-symbolic numerical quantities sets the basis for the development of formal mathematical competence and if numerosity perception is limited by crowding, could stronger than normal crowding mechanisms contribute to the origin of DD? By hampering efficient individual items' identification, abnormal crowding mechanisms might impair ensemble perception leading to the recruitment of texture-density mechanisms even for relatively sparse ensembles. This might then induce a cascade of events: because items might appear too cluttered to be clearly segregated, the representation of numerosity might be imprecise and understanding the meaning of symbolic numbers more difficult, posing challenges to the development of efficient symbolic mathematical skills.

The effect of crowding on the development of numerical cognition has not been investigated so far. Crowding has been classically studied using letter stimuli (Flom, Heath, & Takahashi, 1963; Bouma, 1970; Toet & Levi, 1992) for its effect on reading: we can only read letters within our central "uncrowded" window while letters in the periphery will be undistinguishable. Our reading rate is limited by crowding, it depends on the observer's critical spacing (the smallest distance between items that avoids crowding) and on the spacing between the viewed letters (Legge, Mansfield, & Chung, 2001; Pelli et al., 2007; Yu, Cheung, Legge, & Chung, 2007). There occurs a developmental expansion of the size of this uncrowded window from the 3rd grade to adulthood which parallels reading speed in normal readers (Bondarko & Semenov, 2005; Kwon, Legge, & Dubbels, 2008). Interestingly, there is evidence that some individuals with developmental dyslexia, a specific learning disability affecting reading skills, have increased crowding with respect to proficient readers (Bouma, 1970; Spinelli, De Luca, Judica, and Zoccolotti, 2002; Martelli, Di Filippo, Spinelli, & Zoccolotti, 2009; Moores, Cassim, & Talcott, 2011; Perea, Panadero, Moret-Tatay, & Gómez, 2012; Zorzi et al., 2012; Callens, Whitney, Tops, & Brysbaert, 2013; Moll & Jones, 2013; Cassim, Talcott, & Moores, 2014; Montani, Facoetti, & Zorzi, 2015; for a review see:

Gori and Facoetti, 2015). More recently it has been proposed that crowding might contribute to the origin of developmental dyslexia (Spinelli et al., 2002; Gori & Facoetti, 2015), although it cannot fully account for it: dyslexics' reading rate still remains slower than proficient readers even after compensating for crowding (Martelli et al., 2009).

Given the reviewed evidence for a crucial contribution of crowding-like mechanisms to the processing of visual numerosity, in the current study we investigated whether excessive crowding is present in DD. Importantly, given that DD is often found in comorbidity with developmental dyslexia (Rubinsten, 2009; Wilson et al., 2015), we tested whether abnormal crowding mechanisms are present in DD independently of associated reading disorders. Moreover, given that DD is also often associated with ADHD (Rubinsten, 2009) and considering that one of the most influential model of crowding attributed the occurrence of this phenomenon to limited attentional resources (He, Cavanagh, & Intriligator, 1996), in the current study we also evaluated crowding in DD independently of major visual attentional deficits, potentially indicative of ADHD. To this aim, we measured visual crowding effects in DD individuals and compared their performance against a group of matched controls. In order to investigate visual crowding independently of non-symbolic number comparison abilities, we used stimuli and tasks that were not related to number processing and asked participants to judge the orientation of a Gabor patch shown in isolation or surrounded by flankers. We further measured participants' ability to discriminate sparse and dense non-symbolic ensembles, hypothesising that DD subjects would specifically fail in the former, suggesting a specific fragility of the number system. Finally, we explored the relation between visual crowding, numerical and mathematical abilities, and we speculate on the possible contribution of crowding mechanisms in the development of numerical cognition.

## Material and methods

### Subjects

Seventeen adults without mathematical impairment and seventeen adults with mathematical impairment were included in the study. Participants without mathematical impairment were recruited through a diffusion list provided by the CNRS, primarily directed to cognitive psychology students, but open also to students from other faculties.

Contacts with participants with mathematical impairment were provided either by our speech therapist and neuropsychologist collaborators or

obtained through advertisements on social media and in universities. The advertisement encouraged people with mathematical difficulties to fill in an online screening questionnaire. In addition to collecting general information (such as age and schooling level), the first part of the questionnaire investigated whether the individual had received a diagnosis of dyscalculia or neurological disorders. The second part of the questionnaire explored whether the claimed math difficulties were sufficiently strong to impact the individual's everyday life (for example when dealing with money or quantities in social environments and everyday activities), or to impact the ability to perform some basic numerical tasks (such as counting or reading/writing numerals or solving simple mathematical operations without using fingers or a calculator).

To be included in the experiment, all participants were required to (a) be between 18 and 50 years old, (b) present no neurological disorder, and (c) have completed at least secondary level education. Furthermore, participants that were included in the math impaired group needed to either have been clinically diagnosed with dyscalculia by a neuropsychologist or speech therapist or have claimed major difficulties when dealing with numbers according to the questionnaire. Participants fulfilling these criteria were contacted to participate in an extensive neuropsychological assessment, where measures of verbal and non-verbal intelligence, verbal and visuospatial working memory, visual attention, reading abilities, inhibitory skills and mathematical performance were obtained. In a separate session held in a different day, participants performed a series of psychophysical experiments.

Two subjects included in the math impaired group dropped from the study: one subject who initially showed interest in participating in the study was not available for the proposed testing sessions and not further contactable afterward. The other subject underwent the neuropsychological assessment but was never available for the second testing session requiring participants to perform the psychophysical tests.

To define the final DD and control groups, we performed an additional selection based on the results of the neuropsychological assessment. Specifically, we z-scored the participants' results to the math tests: first, we calculated the mean and standard deviation of the scores obtained by the participants without math difficulty in each test, then we used these values for normalization, i.e. we subtracted the mean of the group without math difficulty from the score of each participant (including DD) and then divided it by the standard deviation of the participants without math difficulty. Mathematical performance was considered below the normal level if z-scores calculated from either accuracy or reaction time in two (of a total of four) or

more math tests exceeded the average z-scores of the non-math-impaired group by more than 2 standard deviations. All DD participants exceeded this cut off. Two participants in the non-math-impaired group exceeded this cut off and where therefore discarded. The same procedure was applied to the accuracy and reading speed of a reading test and of a visual search test (see below for test description) in order to identify DD subjects who also had major associated reading or attentional deficits, potentially reflecting associated dyslexia or ADHD disorders. Two DD subjects exceeded the cut-off for reading abilities (one for number of errors, the other for speed) and other two exceeded the cut-off for the number of errors in the visual search test.

Overall, fifteen participants in the control group (age 31±10, 8 females) and fifteen participants in the DD group (age 27±11, 10 females), four of which with associated reading or attention difficulties, were included in the study. All participants had normal or corrected to normal visual acuity.

Written informed consent was obtained from all participants in accordance with the Declaration of Helsinki, and the study was approved by the research ethics committee of University Paris-Saclay.

## Neuropsychological assessment

Prior to the psychophysical experiments, subjects underwent neuropsychological assessment. We selected the subtests Similarities and Matrix Reasoning from the Wechsler Adult Intelligence Scale IV edition (WAIS-IV) as a measure of verbal and non-verbal IQ. As a measure of verbal working memory, we selected the digit span subtest from WAIS-IV, while the Corsi Block Tapping test was used to measure visuospatial working memory.

Reading abilities were assessed with the "Alouette" (Lefavrais, 1967), the most widely-used French reading test. It involves reading aloud a brief text composed of grammatically plausible sentences, including existing regular and irregular words, without a clear overall meaning. The time needed to read the text and the number of errors made were measured.

To measure inhibitory skills, the Stroop-Victoria test adapted for francophone subjects (Bayard, Erkes, & Moroni, 2009) was administered. Participants were required to pronounce as quickly as possible the color of the ink of a series of filled circles, of a list of words ('mais', 'pour', 'donc', 'quand', meaning 'but', 'for', 'so', 'when') and of a list of color words ('jaune', 'rouge', 'vert', 'bleu', meaning 'yellow', 'red', 'green', 'blue'). For the color words, the color of the ink was always incongruent with the meaning (for example 'rouge', meaning red, written in blue). The interference index was calculated by dividing the time needed to

perform the task with the color words by the time needed to name the color of circles.

As a measure of visual attention, visual search performance was measured with the Bells test. Participants were shown a paper sheet containing black silhouettes of different objects. They were required to identify and cross with a pen all the bells embedded in the sheet. Time recording was stopped when the participant considered all the bells crossed and the number of omitted bells was recorded.

To evaluate mathematical abilities, participants were tested with several subtests of the French battery TEDI Math Grands (Noël & Grégoire, 2015). This computerized battery measures the individual's performance over a range of tests targeting various basic numerical abilities. Subjects were required to: 1) estimate the number of briefly presented dots within the small numerosity range (1–6 items); 2) compare two single-digit Arabic numerals; 3) mentally perform single-digit multiplications and subtractions. The software collected the participants' accuracies and reaction times for most of these tests, with the exception of the test measuring the ability to estimate numerosities for which only the accuracy was recorded.

We calculated standard scores for the IQ subtests (Similarities and Matrix reasoning), for the verbal and visuospatial working memory and for inhibition referring to standardized norms for adults. For the TEDI-MATH we analyzed the number of correct responses and, when recorded, the reaction time (in ms) needed to respond. Given that reaction time and accuracy can often inversely trade off with each other, we reduced the number of variables by calculating the inverse efficiency score (IES, Collins, Park, & Behrmann, 2017), corresponding to the reaction time (RT) divided by the proportion of correct responses. From the TEDI-MATH results, we computed: 1) IES Digits obtained from the results of the single-digit Arabic numeral comparison test; 2) IES Calculation – obtained by averaging together the results from the multiplication and subtraction tests and computing the IES from the combined measure; 3) IES Math – obtained by averaging the IES Digits and IES Calculation as index of general math ability.

Independent sample t-tests were performed to evaluate differences across groups. These tests were applied to either the standardized test scores described (for the IQ, memory and inhibition tests) or to the raw scores in the cases where the norms did not cover the adult age range (in the case of the math, reading and visual search tests).

## Psychophysical experiments

All visual stimuli used in the psychophysical experiment were viewed binocularly from approximately 60 cm in a dimly lit room, displayed on a 15-inch Laptop (HP) LCD monitor with $1600 \times 900$ resolution at refresh rate of 60 Hz. Stimuli were generated and presented under Matlab using PsychToolbox routines (Brainard, 1997).

### Visual crowding: Experiments 1–3

Stimuli were oriented Gaussian-windowed sinusoidal gratings (carrier frequency 1 cycle/visual degree, the gaussian window around the Gabors had a standard deviation of 0.85°, Michelson contrast of 0.9%) presented on a gray background. Stimuli were displayed randomly to the left or right of the central fixation point at 10° of eccentricity (Figure 1). Subjects were asked to maintain central fixation and to judge whether the external half of the grating (with respect to the screen center) was tilted up or down of horizontal by pressing either the up or down arrows on the keyboard. At the beginning of each trial, a central fixation point was shown for 1500 ms, then the stimuli were briefly presented for 33 ms. Participants were instructed to provide their response in the shortest time possible. In Experiment 1, the oriented grating (Figure 1A) was presented in isolation, whereas in Experiment 2 (Figure 1B), the target grating was presented simultaneously with two other gratings (identical features, but with fixed horizontal orientation) that served as distractors (flankers). The two flankers vertically surrounded the target, at a fixed and very close center-center distance (0.85°). In Experiment 2, participants were asked to report the orientation of the central grating (target), ignoring the two flankers.

In both Experiment 1 and 2, the target orientation was adaptively changed according to the subject's responses following a QUEST algorithm (Watson & Pelli, 1983). Participants performed one session of 60 trials for each experiment. Responses faster than 150 ms or longer than two standard deviations with respect to each participant's average reaction time in each condition were discarded from the analysis. For each participant, the proportion of correct responses against the target orientation angle was fitted by a cumulative Gaussian function. The orientation angle needed for the subject to score 75% correct defined the participant's orientation threshold. To verify the reliability of the estimated thresholds, we calculated the goodness of fit for each subject and experiment and set a minimum limit at $R^2 = 0.75$. The quality of the fit did not meet this criterion for the data obtained from one DD participants in Experiment 1 and two DD participants in Experiment 2, confirming their verbally reported difficulty in performing these tasks. Given that a reliable estimate of the orientation discrimination threshold in the condition tested in Experiment 2 was needed to perform Experiment 3 (see below), data from these three subjects was discarded from the analyses of
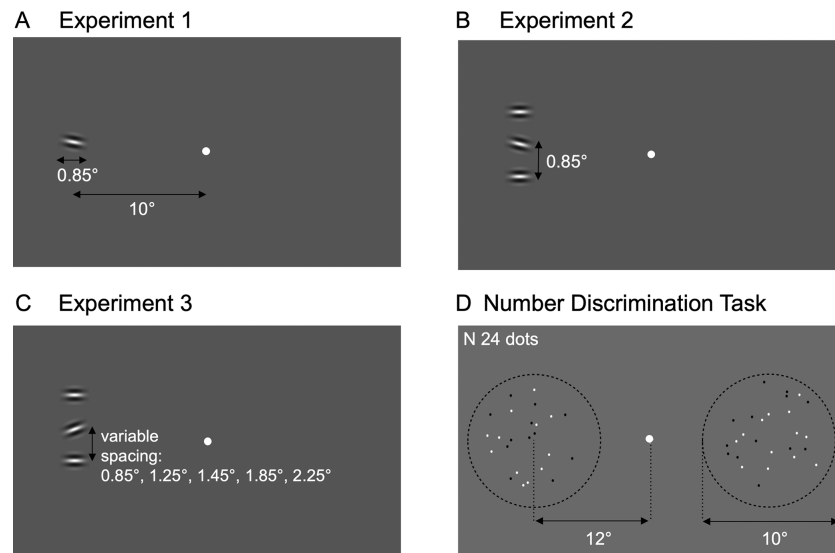
Figure 1. Examples of stimuli used in the visual crowding and number discrimination experiments. (A) In Experiment 1, a single oriented Gabor grating was briefly shown (33ms) at 10° of eccentricity either to the left or right of the central fixation point. Participants were asked to judge the grating's orientation with respect to the imaginary horizontal axis and to press the up or down arrow key (the correct response in this trial is 'up'). (B) In Experiment 2, three oriented gratings were shown at a very close distance and the participants' task was to judge the orientation of the central one, ignoring the two flankers. (C) Stimuli and task in Experiment 3 were the same as in Experiment 2, but the target-flanker distance was varied, to modulate the strength of the crowding effect (the larger the distance, the weaker the crowding effect). (D) In the number discrimination task, two arrays of dots were briefly (250 ms) and simultaneously presented. Subjects were asked to indicate which side of the screen contained more dots.

Experiment 1–3. Orientation discrimination thresholds and reaction times measured in Experiment 1 and 2 were entered into a repeated measures ANOVA with crowding condition (2 levels: uncrowded/crowded) and group (2 levels: control/DD) as within- and between-subject factors, respectively.

Similar to Experiment 2, in Experiment 3 participants were presented with three vertically positioned oriented gratings and asked to evaluate the orientation of the central target (Figure 1C). The central target orientation was fixed at the participant's orientation discrimination threshold, estimated from Experiment 2. The target-flanker distance was varied at every trial to measure the spatial extent of crowding (i.e. the critical spacing). Five target-flanker distances were tested: 0.85° (same as Experiment 2), 1.25°, 1.45°, 1.85° and 2.25°. Trials in Experiment 3 were presented using the method of constant stimuli: participants performed 20 trials for each of the 5 target-flanker distances and 2 presentation sides (left and right of the fixation point), for a total of 200 trials. As for Experiment 1 and 2, responses faster than 150 ms or longer than two standard deviations with respect to each participant's average reaction time were discarded from the analysis. Response accuracies were entered into a repeated measures ANOVA with target-flanker distance (5 levels: the 5 target-flanker distances) and group (2 levels: control/DD) as the within- and between-subject factors

respectively. Furthermore, we computed the critical spacing, defined as the target-flanker distance at which the response accuracy reached the plateau value of 90%, following the procedure used by (Freyberg, Robertson, & Baron-Cohen, 2016). To this end, we plotted the accuracy as a function of the target-flanker distance and fitted the data with an exponential curve using the following equation:

$$Y = a * \exp(b * x) + c$$

where x refers to the target-flanker distance, with the constraints that the parameters a and b had to be negative, and c had to fall between 0.5 and 1. As starting values for the parameters a, b, and c we chose $-0.2$, $-2$, and $0.95$ respectively. This ensured that the percent accuracy increased with target-flanker distance and plateaued to a value between 50% and 100% accuracy. The critical spacing was therefore calculated as:

$$\text{critical spacing} = \ln(-0.1 * c/a)/b$$

and compared across groups by means of independent sample t-test.

Reaction times were also recorded and combined with response accuracies to calculate the inverse efficacy score (IES). Reaction times and IES were entered

into ANOVA with target-flanker distance (5 levels: the 5 target-flanker distances) and group (2 levels: control/DD) as the within- and between-subject factors, respectively.

Similarly to (Cassim et al., 2014), we calculated a crowding index defined as the difference between IES at the smallest (0.85°) and the largest (2.25°) target-flanker distance. This measure reflects the accuracy-normalized reaction time cost of performing the orientation discrimination task in the stronger crowding compared to weaker crowding condition.

The statistical analyses which resulted significant when testing the whole group of DD subjects who successfully performed Experiment 1–3 (twelve subjects), were replicated using a reduced group which did not include the four DD subjects who had associated reading or attentional difficulties.

For all the statistical analyses mentioned in this section we additionally performed Bayesian analysis using JASP (JASP Team, 2020). Hypotheses were always tested two-sided and using the default prior distribution, as recommended in case where there is no a priori knowledge of the strength of the effects (van Doorn et al., 2019; Faulkenberry, Ly, & Wagenmakers, 2020). For Bayesian ANOVA, models including each individual factor, both factors, and both factors plus their interaction were ordered by their predictive performance relative to the best model. The reported Bayes Factors (BF) for main effects and interaction terms correspond to inclusion Bayes factors resulting from the analysis of effects across "all matched models" in JASP (van den Bergh et al., 2019). For the interaction effect, this is equivalent to the BF resulting from comparison of two models including both main factors in each case and either including or excluding the interaction term. Bayes Factors are always reported alongside with classical statistics, however, to help readability, results are discussed in Bayesian terms only when the classical analyses suggested that there was a difference between groups. Bayes factors should be interpreted as providing positive (BF = 1–3), substantial (BF = 3–10), strong (BF = 10–30), very strong (BF = 30–100) and decisive (BF > 100) evidence in favor of the alternative hypothesis. BFs below 1 lend support the null hypothesis. Specifically, Bayes factors should be interpreted as providing positive (BF = 0.3–1), substantial (BF = 0.1–0.3), strong (BF = 0.03–0.1), very strong (BF = 0.01–0.03) and decisive (BF<0.01) evidence in favor of the null hypothesis.

### Non-symbolic number discrimination task

Stimuli consisted of array of dots, half white and half black, presented on a mid-gray background, so that luminance was not a cue for number. Two arrays of dots were simultaneously and briefly (250 ms) presented on either side of the central fixation point, at 12° of

eccentricity (Figure 1D). Participants performed a comparison task indicating the side of the screen with the more numerous array. In separate sessions, the numerosity of the probe array (randomly shown to the left or to the right of the fixation point) was fixed at 24 (N24, sparse array) or at 64 (N64, dense array) dots, while the numerosity of the test array adaptively changed, according to the participant's responses, following the QUEST algorithm. Dots were 0.1° diameter large and constrained to fall within a virtual circle of 10° diameter (yielding a density of 0.3 and 0.82 dots/deg$^2$ for N24 and N64, respectively). The minimal inter-dot distance was 0.3°. Participants performed two separate sessions of 35 trials each, testing the two numerosities, with half the participants starting with N24 and the other half with N64. The proportion of trials where the test appeared more numerous than the probe was plotted against the logarithm of the test numerosities and fitted with a cumulative Gaussian function. The mean of this function provided an estimate of the point of subjective equality (PSE, a measure of perceived numerosity), while the standard deviation was used to estimate the precision (i.e., the just-noticeable difference, JND), which was divided by the point of subjective equality to estimate the Weber fraction. Weber fractions were entered into an ANOVA with numerosity (2 levels: N24 and N64) and group (2 levels: control/DD) as within and between subjects' factors respectively. All fifteen DD subjects successfully performed the experiment and were entered in the analysis. Bayesian analyses were additionally performed as specified in the previous section.

## Results

### Neuropsychological assessment

In the interview conducted during the neuropsychological assessment, all participants confirmed being free of neurological disorders and having had access to appropriate education during school-age. Among the participants included in the DD group, four had received formal diagnosis of dyscalculia during childhood and the others confirmed having always had difficulties whenever dealing with numbers and quantities and major problems in acquiring mathematical skills since the early school years. All subjects claimed that the mathematical difficulties persisted over years. Furthermore, 7 participants out of 15 reported having at least one relative with difficulty in mathematics, reading, writing, or orthography.

The DD and control group were not significantly different in age, verbal and non-verbal IQ, reading accuracy, inhibitory control, as measured by the

| | Control group (N = 15) mean (STD) | Dyscalculic group (N = 15) mean (STD) | Statistical analysis t-value |
|---|---|---|---|
| Age | 31 (10) | 27 (11) | −0.93 |
| IQ | | | |
|   Similarities | 13 (3) | 13 (2) | 0.42 |
|   Matrices | 12 (3) | 10 (3) | −1.83 |
| Reading Ability | | | |
|   Time (seconds) | 89 (14) | 106 (22) | 2.47* |
|   N errors | 3 (3) | 4 (3) | 1.43 |
| Working memory | | | |
|   Verbal (Digit span) | 12 (3) | 9 (3) | −2.59* |
|   Visuospatial (Corsi) | 13 (2) | 10 (2) | −3.27** |
| Inhibition | | | |
|   Color Stroop Score | 12 (2) | 11 (4) | −1.15 |
| Visuo-Spatial attention | | | |
|   Time (seconds) | 105 (45) | 112 (37) | 0.51 |
|   N omissions | 1 (1) | 2 (3) | 0.53 |
| Numerical skills/Arithmetics TEDI-MATH (no of items) | | | |
|   Small numerosity estimation (36) | 33 (4) | 32 (3) | −0.75 |
|   Digit Comparison | | | |
|     Accuracy (48) | 46 (1) | 47 (2) | 1.04 |
|     Reaction Time (ms) | 558 (51) | 739(168) | 3.97** |
|     IES Digit (ms) | 578 (54) | 754 (159) | 4.03** |
|   Multiplication | | | |
|     Accuracy (20) | 18 (2) | 15 (2) | −5.05** |
|     Reaction Time (ms) | 1681 (403) | 3617 (1681) | 4.34** |
|   Subtraction | | | |
|     Accuracy (20) | 19 (1) | 18 (2) | −2.18* |
|     Reaction Time (ms) | 1572 (333) | 3307 (1362) | 4.79** |
|   Calculation (x and -) | | | |
|     IES Calculation (ms) | 3481 (727) | 8365 (3559) | 5.20** |
|     IES General Math (ms) | 1166 (201) | 2007 (1156) | 2.77** |

Table 1. DD differs significantly from controls at: *p < 0.05, **p < 0.01.

Color-Stroop test, and visual search performance, as measured by the Bells test, (all p-values > 0.05, see Table 1 for descriptive statistics and tests across groups). The DD and control groups significantly differed in reading speed ($t_{(27)} = 2.47$, $p = 0.02$), verbal ($t_{(28)} = -2.59$, $p = 0.01$) and visuo-spatial working memory ($t_{(28)} = -3.27$, $p = 0.002$), and most of the numerical and arithmetical tests. Specifically, the DD group was slower when comparing digits ($t_{(28)} = 3.97$, $p = 0.0004$), performing mental multiplication ($t_{(28)} = 4.34$, $p = 0.0002$) and subtraction ($t_{(28)} = 4.79$, $p = 0.00005$). Accuracy for mental multiplication and subtraction was also significantly lower with respect to the control group (for multiplication: $t_{(28)} = -5.05$, $p = 0.00002$; for subtraction $t_{(28)} = -2.18$, $p = 0.04$). IES for digit comparison ($t_{(28)} = 4.03$, $p = 0.0004$), calculation ($t_{(28)} = 5.20$, $p = 0.00002$) and general math ($t_{(28)} = 2.77$, $p = 0.009$) significantly differed across groups.

## Visual crowding

One participant from the DD group did not succeed in completing Experiment 1, which required to judge the orientation of one isolated grating. The participant claimed that the difficulty was not due to discriminating the orientation of each individual grating, but to associating a pair of orientations (each one the mirror image of the other with respect to a vertical imaginary symmetry axis) with the corresponding response (either up or down arrows). To check that the participant did not have a general orientation discrimination impairment we measured the ability to discriminate the orientation of a central grating around +45 or −45 degrees (for details on the stimuli and procedure see: Castaldi, Tinelli, Cicchini, & Morrone, 2018a). The participant showed very good thresholds for both orientations (3 and 1 visual degrees for +45 and −45 degrees respectively), comparable to typical subjects'
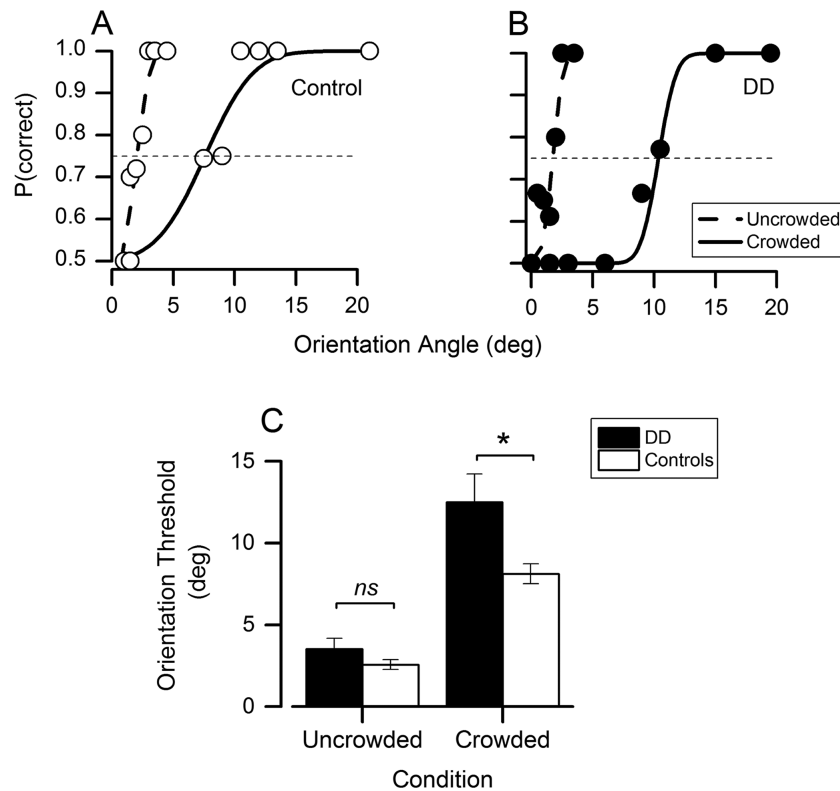
Figure 2. Effect of visual crowding on orientation discrimination thresholds. Psychometric functions for two subjects from the control (A) and the DD (B) group. The proportion of correct responses is plotted as a function of the orientation angle with respect to the horizontal (in visual degrees) for the uncrowded (hatched curve, Experiment 1) and crowded (solid curve, Experiment 2) conditions. Orientation thresholds were measured at the point in which the proportion of correct responses reached 75% (dashed lines). (C) Average orientation discrimination thresholds in the DD (black bars) and control (white bars) groups for the two crowding conditions. The addition of flankers affected orientation discrimination performance in both groups, but this effect was stronger in the DD group with respect to the control group.

average discrimination threshold of 4.4±0.7 (typical values taken from: Castaldi et al., 2018a). Two other subjects from the DD group, while easily performing Experiment 1, failed to complete Experiment 2, in which the target grating was surrounded by two flankers placed in close proximity. The participants claimed that the task was too difficult because they could not distinguish the central target from the flankers. These qualitative verbal reports were confirmed by the poor data quality which did not allow the fitting to provide reliable thresholds (goodness of fit $R^2$ lower than 0.15 for both subjects, see methods). As a consequence, it was not possible to test these three subjects with Experiment 3 for which a measure of the participant's orientation threshold at the closest target-flanker distance (estimated in Experiment 2) was needed. These subjects were therefore excluded also from Experiment 1 and the results described below were obtained from the remaining twelve subjects included in the DD group.

Results from Experiment 1 and 2 are shown in Figure 2. Figure 2A shows the psychometric curves of two representative single subjects from the control

(Figure 2A) and DD (Figure 2B) group in Experiment 1 (uncrowded condition, hatched curves) and Experiment 2 (crowded condition, solid curves). The rightward shift of the solid psychometric curves on the x-axis indicates that the orientation discrimination thresholds increased under crowding for both groups. The orientation discrimination thresholds measured in the two groups and crowding conditions are shown in Figure 2C. Orientation discrimination thresholds were similar across groups in the uncrowded condition (2.57°±1.19° for the control group and 3.53°±2.25° for the DD group) and increased in the crowded condition in both groups. However, the presence of flankers surrounding the target induced a much larger increase in the orientation discrimination thresholds in the DD group with respect to the control group (8.12° ± 2.37° for the control group and 12.48° ± 5.96° for the DD group). To statistically test for these differences, we entered the orientation discrimination thresholds into a repeated measures ANOVA with crowding condition (2 levels: uncrowded/crowded) and group (2 levels: DD/controls) as within- and between-subject factors, respectively. The

significant interaction between crowding condition and group ($F_{(1,25)} = 5.01$, p = 0.03, $\eta p^2 = 0.16$, BF = 2.3) confirmed that although the orientation discrimination thresholds under crowding significantly increased both in the DD ($t_{(11)} = -5.89$, p<$10^{-5}$, Cohen's $d = 1.69$, BF = 280.26) and in the control group ($t_{(14)} = -8.97$, p<$10^{-5}$, Cohen's $d = 2.32$, BF = 45110.74) with respect to the uncrowded condition, this increase was stronger for the DD group with respect to the control group. Post hoc tests showed that discrimination thresholds were significantly higher in the DD group compared to the control group under crowding ($t_{(25)} = -2.59$, p = 0.016, Cohen's $d = 0.96$, BF = 3.67). Importantly, this difference could not be attributed to an overall poorer orientation discrimination ability in DD, as the orientation discrimination thresholds in the uncrowded condition were not statistically different across groups ($t_{(25)} = -1.41$, p = 0.17, Cohen's $d = 0.55$, BF = 0.72). In Bayesian terms, the above BFs provide positive evidence for including the interaction between crowding condition and group as a predictor of orientation thresholds. There was also substantial evidence in favor of the hypothesis that under crowding the orientation discrimination thresholds were higher in the DD group compared to the control group.

Reaction times of the DD group were comparable to the control group in the uncrowded condition (RTs control group = 0.81±0.15, RTs DD group = 0.97±0.25) and visual crowding slowed down reaction times in both groups (RTs control group = 0.93±0.21, RTs DD group = 1.05± 0.17). Repeated measure ANOVA revealed a significant main effect of crowding condition ($F_{(1, 25)} = 8.97$, p = 0.006, $\eta p2 = 0.26$, BF = 8.51), but no significant main effect of group ($F_{(1,25)} = 3.88$, p = 0.06, $\eta p2 = 0.13$, BF = 1.55), nor significant interaction between group and crowding condition ($F_{(1, 25)} = 0.15$, p = 0.70, $\eta p2 = 0.006$, BF = 0.40), suggesting that visual crowding slowed down reaction times in the two groups to the same extent.

In sum, orientation discrimination thresholds in peripheral vision were comparable across groups when stimuli were presented in isolation, and visual crowding affected performance in both groups, as reflected by both the higher orientation discrimination threshold and the longer reaction times in the crowding condition. However, visual crowding caused a stronger increase in the orientation discrimination thresholds in the DD group with respect to the control group.

In Experiment 3 we measured the spatial extent of the visual crowding effect in the two groups, by fixing the target orientation at each individual subject's threshold, measured in Experiment 2, and varying the target-flanker distance. The orientation discrimination performance was expected to progressively improve with larger target-flanker distances, as the flankers should progressively fall outside the crowding window and stop

interfering with the central target perception. Figure 3A shows that indeed the proportion of correct responses increased with larger target-flanker distances in both groups. The proportion of correct responses was entered into a repeated-measures ANOVA, with target-flanker distance (5 levels) and group (2 levels) as the within- and between-subject factors, respectively. The proportion of correct responses started at 0.75 for the shortest target-flanker distance (as expected given that this distance corresponded to the one at which the orientation threshold was estimated in Experiment 2), and increased with larger target-flanker distances in both groups (significant main effect of target-flanker distance: $F_{(4,44)} = 53.20$, p<$10^{-5}$, $\eta p^2 = 0.89$, BF = $3.6*10^{20}$), up to 0.9. The magnitude of this effect was comparable across the two groups (no significant main effect of group: $F_{(1,25)} = 0.14$, p = 0.71, $\eta p^2 = 0.01$, BF = 0.46; no interaction between group and target–flanker distance: $F_{(4,44)} = 1.97$, p = 0.12, $\eta p^2 = 0.14$, BF = 1.58). The critical spacing, although on average slightly higher in the DD with respect to the control group (DD group: 1.78°±0.88°, control group: 1.57°±0.79°), was not significantly different ($t_{(25)} = 0.64$ p = 0.52, Cohen's $d = 0.25$, BF = 0.41).

Reaction times (Figure 3B) and IES (Figure 3C) decreased as a function of the target-flanker distance, meaning that responses speeded up and the orientation discrimination performance improved as the flankers were displayed farther away from the target in both groups.

A repeated measures ANOVA confirmed that reaction times significantly decreased with target-flankers distance in both groups (significant main effect of target-flanker distance: $F_{(4,100)} = 8.95$, p<$10^{-5}$, $\eta p^2 = 0.26$, BF = 1169.85), however the DD group was overall slower with respect to the control group (significant main effect of group: $F_{(1,25)} = 5.31$, p = 0.03, $\eta p^2 = 0.18$, BF = 1.82, no significant interaction between group and target–flanker distance: $F_{(4,100)} = 2.21$, p = 0.07, $\eta p^2 = 0.08$, BF = 1.05).

These results suggest that in order to perform the orientation discrimination task with response accuracy comparable to the control group, the DD group needed more time. The IES decreased with target-flanker distance at a different rate across groups, as shown by the significant interaction between target-flanker distance and group ($F_{(4,100)} = 4.02$, p = 0.005, $\eta p^2 = 0.13$, BF = 7.33). With respect to the control group, the DD group showed a particularly poor performance in target orientation discrimination (higher IES, reflecting the combination of lower response accuracies and longer response times) when the flankers were shown at the closest distances from the target: post-hoc tests showed that the difference in IES across groups was statistically significant for the target-flanker distances of 0.85° (p = 0.005, BF = 8.50), 1.25° (p = 0.03,
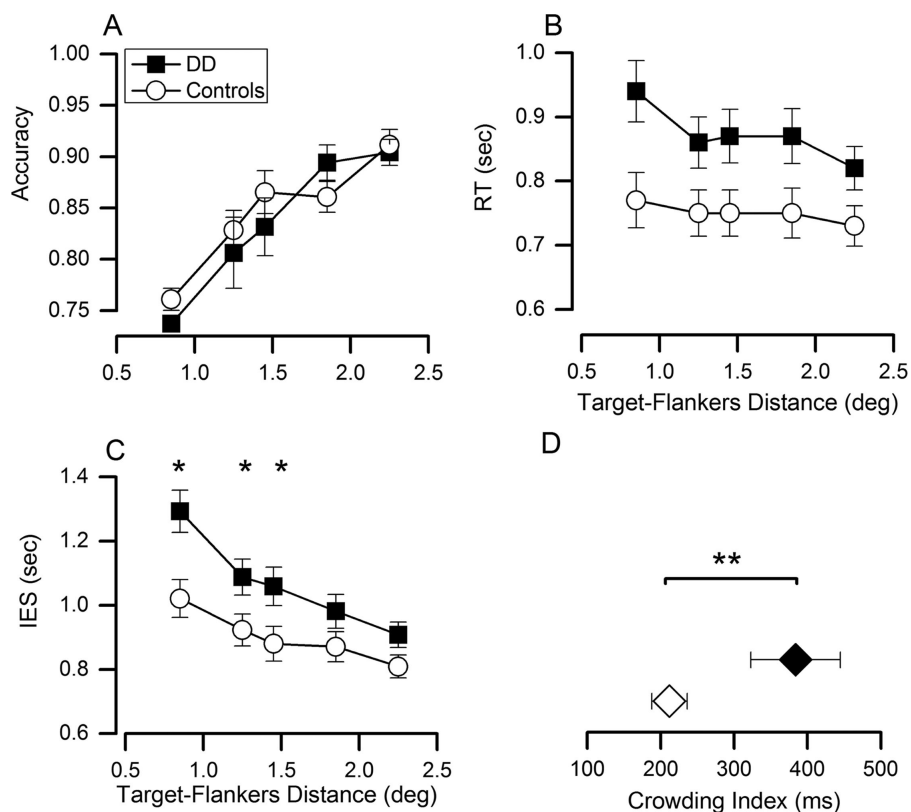
Figure 3. Spatial extent of the visual crowding effect. (A) The proportion of correct responses is plotted as a function of the target-flanker distance for the DD (black symbols) and the control (white symbols) group. As expected, response accuracy increased with larger target-flanker distance. Reaction times (B) and IES (C) decreased with larger target-flanker distance in both groups. IES was significantly higher in DD with respect to the control group, especially for the closest target-flanker distances. The crowding index (D) was higher in the DD with respect to the control group.

BF = 2.00) and 1.45° (p = 0.03, BF = 1.96). With larger target-flanker distances the performance across groups became progressively more comparable, and the IES differences were not significantly different for target-flanker distances of 1.85° (p = 0.13, BF = 0.84) and of 2.25° (p = 0.08, BF = 1.17). In Bayesian terms, the above BFs provided substantial evidence for including the interaction between target-flanker distance and group as predictor of IES. Moreover, there was substantial evidence in favor of the hypothesis that IES was larger in the DD group compared to the control group for the closest target-flanker spacing and positive evidence that this effect also extended to some of the other distances. The crowding index (Figure 3D), reflecting the difference in IES between the displays with the smallest and the largest target-flanker distance, was significantly higher in DD compared to the control (DD group: 384.63 ms ± 213.02 ms, control group: 212 ms ± 95.02 ms, $t_{(25)}$ = 2.81, p = 0.009, Cohen's d = 1.04, BF = 5.39). The BF points at a substantial evidence in support to this hypothesis.

In sum, the results of Experiment 3 showed that the enhanced orientation discrimination difficulty in

the DD group persisted over several distances between target and flankers compared to controls, as indicated by combined accuracy and reaction time measures.

In order to test for the possibility that the across group differences observed in the orientation discrimination thresholds under crowding (Experiment 2) and in the spatial extent of visual crowding (Experiment 3) were driven by the results of participants having associated reading difficulties or attentional deficits, we repeated the analysis of Experiment 1–3 on a reduced group of eight DD subjects, after excluding those who scored more than 2 standard deviations distance from the control group's mean in speed or accuracy in the reading or visual search tests.

For the orientation discrimination thresholds measured in Experiment 1 and 2, the interaction between crowding condition and group was significant ($F_{(1,21)}$ = 4.92, p = 0.03, $\eta p^2$ = 0.19, BF = 2.2) showing that the increase in orientation discrimination thresholds under the crowded with respect to the uncrowded condition, observed both for the DD ($t_{(7)}$ = −4.35, p = 0.003, Cohen's d = 1.54, BF = 15.90) and for the control group ($t_{(14)}$ = −9.04, p<10$^{-5}$, Cohen's

d $= 2.32$, BF $= 45510.74$), was stronger for the DD group with respect to the control group. Across group differences in orientation discrimination thresholds in the uncrowded condition (Experiment 1), remained not significant ($t_{(21)} = -1.20$, p $= 0.24$, Cohen's d $= 0.53$, BF $= 0.63$). Under crowding the DD group showed significantly higher thresholds compared to the control group ($t_{(21)} = 2.41$, p $= 0.02$, Cohen's $d = 1.05$, BF $= 2.65$). In Bayesian terms, there was a positive evidence for the interaction between crowding conditions and group as predictor of orientation thresholds as well as in support of the hypothesis that under crowding the orientation discrimination thresholds were higher in the DD group compared to the control group.

Reaction times measured in Experiment 3 significantly decreased with target-flanker distance in both groups (significant main effect of target-flanker distance: $F_{(4,84)} = 10.35$, p$<10^{-5}$, $\eta p^2 = 0.33$, BF $= 280.66$), however the DD group was overall slower with respect to the control group (significant main effect of group: $F_{(1,25)} = 9.10$, p $= 0.007$, $\eta p^2 = 0.30$, BF $= 5.27$, significant interaction between group and target–flanker distance: $F_{(4,84)} = 3.08$, p $= 0.007$, $\eta p^2 = 0.15$, BF $= 3.94$, all post-hoc p-values $<0.05$).

The IES showed a significant interaction between group and target-flanker distance ($F_{(4,84)} = 4.88$, p $= 0.01$, $\eta p^2 = 0.18$, BF $= 14.87$). The IES difference across groups tended to attenuate for larger target-flanker distances, and was significant for most of them - i.e. for target-flankers distances of 0.85° (p $= 0.001$, BF $= 44.02$), 1.25° (p $= 0.01$, BF $= 5.34$), 1.45° (p $= 0.005$, BF $= 8.47$), 1.85° (p $= 0.03$, BF $= 2.12$), but not for 2.25° (p $= 0.06$, BF $= 1.54$). In Bayesian terms, the above BFs provided strong evidence for including the interaction between target-flanker distance and group as predictor of IES. Moreover, there was very strong evidence in favor of the hypothesis that IES was larger in the DD group compared to the control group for the closest target-flanker distance and substantial to positive evidence that this effect also extended to the other distances.

Finally, the crowding index was significantly higher in the DD group with respect to the control group (DD group: 430.58 ms $\pm$ 212.30 ms, control group: 212 ms $\pm$ 95.02 ms, $t_{(25)} = 3.44$ p $= 0.002$, Cohen's $d = 1.32$, BF $= 14.72$). The BF points at a strong evidence in support to this hypothesis.

Overall, the analyses performed when excluding from the DD group the four participants with reading and attentional difficulties confirmed the previous results.

## Number discrimination

All participants performed two number discrimination tasks in which they were asked to compare the numerosity of two sets of dots.
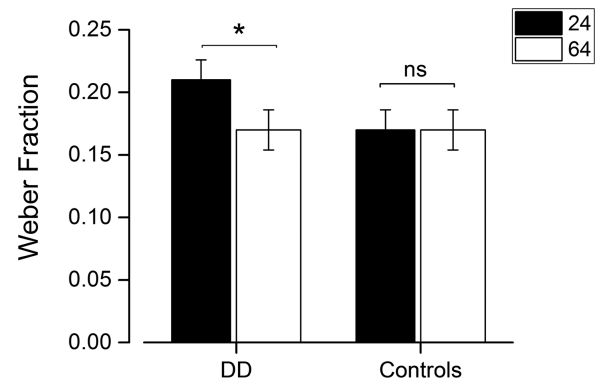


Figure 4. Numerosity Discrimination Task. Average Weber Fraction in the DD and control groups for the two numerosity conditions.

The DD group was on average less precise with respect to the control group when they had to compare relatively sparse arrays (weber fraction for N24 in the control group $= 0.17\pm0.06$, in the DD group $= 0.22\pm0.06$), but not when they had to compare cluttered sets (weber fraction for N64 in the control group $= 0.17\pm0.05$, in the DD group $= 0.17\pm0.06$, Figure 4). To statistically test for these differences, we entered the weber fractions into a repeated measures ANOVA with numerosity (2 levels: N24 or N64) and group (2 levels: DD/controls) as within- and between-subject factors, respectively. The ANOVA revealed a significant main effect of numerosity ($F_{(1,28)} = 4.20$, p $= 0.05$, $\eta p2 = 0.13$, BF $= 1.46$), but the main effect of group ($F_{(1,28)} = 1.60$, p $= 0.21$ $\eta p2 = 0.05$, BF $= 0.62$), and interaction between group and numerosity ($F_{(1,28)} = 2.61$, p $= 0.11$, $\eta p2 = 0.08$, BF $= 1.08$) were non-significant, suggesting that the two groups were overall less precise in comparing sparser with respect to denser arrays of dots. This latter finding is mainly driven by the DD group that showed a statistically larger weber fraction for comparing sparse (N24) compared to dense (N64) arrays ($t_{(14)} = 2.36$ p $= 0.03$, BF $= 2.42$), while in the control group this difference was not observed ($t_{(14)} = 0.34$ p $= 0.73$, BF $= 0.27$). The difference in weber fraction between groups was close to significance for the sparse arrays ($t_{(28)} = 1.90$ p $= 0.06$, BF $= 1.33$) and not significant for the dense arrays ($t_{(28)} = 0.17$ p $= 0.86$, BF $= 0.34$).

## Correlation analyses

In order to explore the relation between visual crowding effects, numerosity perception and mathematical abilities, we performed correlation analyses based on Pearson correlation and Bayesian correlation (we report values for both, but results are discussed in Bayesian terms). We correlated
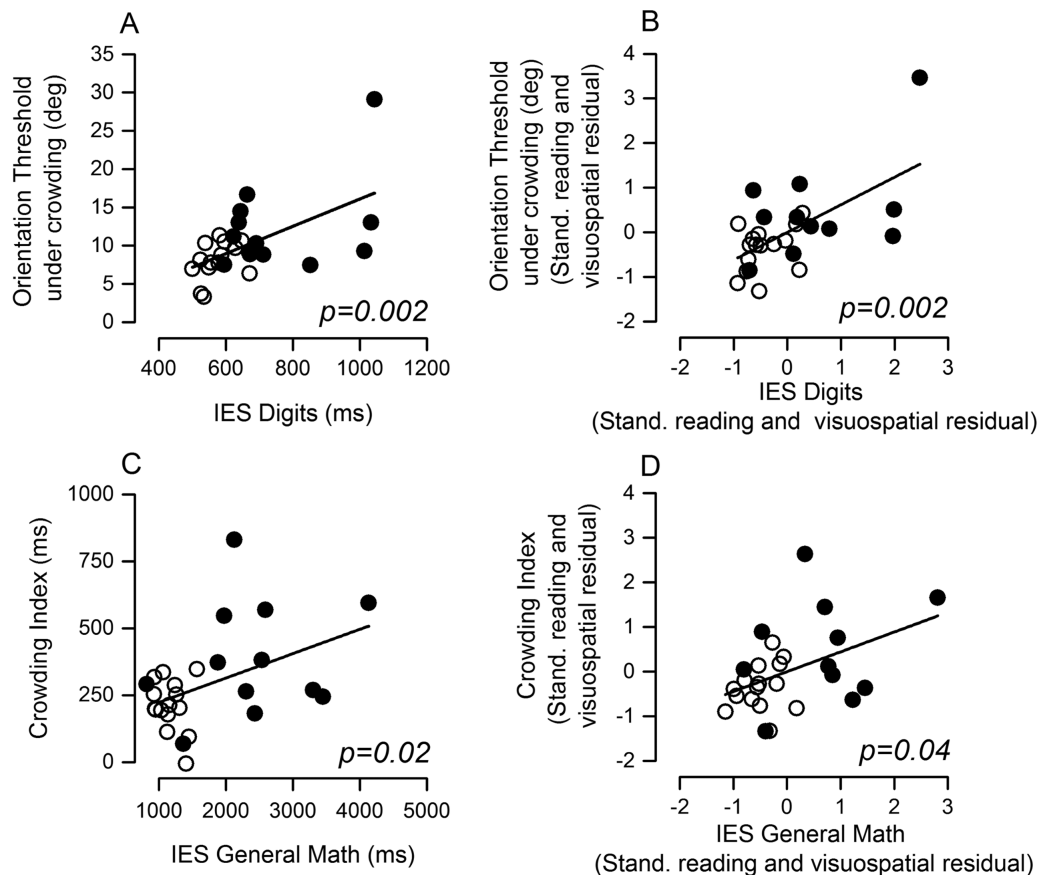
Figure 5. Correlation analysis. Correlation analyses between numerical indices and crowding measures and between their standardized corrected residuals. Orientation discrimination threshold under crowding plotted as a function of IES for digit comparison (A) and plot of the standardized residuals when performance in the reading and visual search tests (B) were regressed out. Same plot but for the crowding index as measured in Experiment 3 (C), and for the standardized residuals (D). White symbols represent control participants, black symbols represent DD participants.

measures of crowding that resulted different between groups with numerosity perception and mathematical abilities. Specifically we correlated: 1) the orientation discrimination thresholds under crowding as measured in Experiment 2, 2) the crowding index calculated from Experiment 3, 3) the weber fraction for N24, 4) the weber fraction for N64, 5) the IES score for digit comparison, 6) the IES score for calculation and 7) the IES score for general math ability. Given that some of the DD subjects presented reading and visuo-spatial attention difficulties we performed the correlation analysis with and without regressing out accuracy and speed measured both in the reading and in the visual search tests.

We found strong evidence in favor of a correlation between orientation discrimination thresholds under crowding and IES for digit comparison ($r_{(27)} = 0.56$, p = 0.002, BF = 20.61, Figure 5A) but not with IES for calculation ($r_{(27)} = 0.27$, p = 0.17, BF = 0.49) nor with IES for general math ($r_{(27)} = 0.28$, p = 0.14, BF = 0.64), nor with the weber fraction at N24 ($r_{(27)} =$

$-0.14$, p = 0.46, BF = 0.25, Suppl. Figure 1A) or N64 ($r_{(27)} = -0.03$, p = 0.85, BF = 0.25, Suppl. Figure 1B). Moreover, there was very strong evidence in favor of correlation between orientation discrimination thresholds under crowding and IES for digit comparison even after controlling for reading and visual search abilities ($r_{(20)} = 0.60$, p = 0.002, Figure 5B, BF = 50.86).

There was substantial/positive evidence in favor of correlation between the crowding index with IES for digit comparison ($r_{(27)} = 0.45$, p = 0.02, BF = 3.33), with IES for calculation ($r_{(27)} = 0.42$, p = 0.03, BF = 2.38) and with IES for general math ($r_{(27)} = 0.45$, p = 0.02, BF = 3.20, Figure 5C), but not with the weber fraction for N24 ($r_{(27)} = 0.09$, p = 0.64, BF = 0.43, Suppl. Figure 1C) or N64 ($r_{(27)} = 0.28$, p = 0.16, BF = 0.25, Suppl. Figure 1D). Moreover, there was strong/positive evidence in favor of a correlation between the crowding index and IES for digit comparison, for general math and for calculation after controlling for reading and visuo-spatial attention abilities ($r_{(20)} = 0.54$, p = 0.009, BF = 11.52; $r_{(20)} =$

0.44, p = 0.04, BF = 2.82, Figure 5D; r$_{(20)}$ = 0.41, p = 0.05, BF = 1.03).

## Discussion

The aim of the present study was to test whether visual crowding mechanisms are altered in individuals with DD independently of major reading and attentional deficits and whether such an impairment relates to the numerical or arithmetical difficulties. Two groups of participants with and without DD were tested with an orientation discrimination task in different crowding conditions. The orientation discrimination thresholds were comparable across groups when the target grating was presented in isolation. When the target grating was surrounded by flankers, the orientation discrimination thresholds increased in both groups, suggesting that they were both subject to crowding effects. Importantly however, the increase in orientation discrimination threshold under crowding was much higher in the DD group with respect to the control group, pointing at stronger crowding effects.

Increasing the target-flanker distance mitigated the detrimental effect of crowding on the orientation discrimination accuracy in both groups to a comparable extent, yet to perform the task, the DD group needed much longer response time at all the target-flanker distances tested with respect to the control group. This difference could not be accounted for by a general tendency to provide slower responses in the DD group, given that reaction times were comparable across groups when participants were required to perform the orientation discrimination task on the isolated target. Evaluation of participants' performance in light of the trade-off between accuracy and reaction time, revealed that the excessive crowding effects observed in the DD group compared to control group extended over several larger target-flanker distances. Overall these results point at the presence of enhanced crowding effects in DD that span over a larger than normal spatial extent.

Importantly, the poorer orientation discrimination performance under crowding with respect to the control group was observed even when removing participants with reading or attentional difficulties from the DD group, suggesting that excessive crowding can characterize DD, independently of dyslexia and attentional disorders.

Several models explaining visual crowding phenomena have been proposed (for a recent review see: Manassi and Whitney, 2018). Perceptual failure under crowding has been often attributed to excessive feature integration (Pelli et al., 2004) or information pooling over a large integration area (Parkes, Lund, Angelucci, Solomon, & Morgan, 2001). According to this account, crowding occurs when integrating the output of

multiple features detectors: target and flankers would fall within the same 'integration field' and thus be perceived as jumbled together. In this framework, crowding effects are generally assumed to occur at an early, pre-attentive stage of visual perception. Opposite to this view, the attentional resolution model of crowding proposed that crowding is due to poor resolution of attention (He et al., 1996). According to this account, spatial attention cannot be directed to individual items if they fall within the minimal selection region of attention, in which case items are perceived as a group, preventing their individual identification (Intriligator & Cavanagh, 2001). In line with this theory, some studies found that directing attention to the target location by pre-cueing it improved identification accuracy, reduced critical distance and recognition contrast threshold (Strasburger, 2005; Yeshurun & Rashal, 2010). More recently, the hierarchical sparse selection model proposed that crowding is not due to degraded sensory representations, but to impoverished sampling of such representations by perception (Chaney, Fischer, & Whitney, 2014). In this view, which can successfully explain why crowding occurs at multiple levels of visual analysis, the limiting factor determining crowding would not be the minimal area of the visual field over which attention can operate, but the sparsity of representation sampling within that region.

In light of these models, DD might be associated with either larger integration fields, coarser attentional resolution or reduced ability to sample information. Regardless of the underlying mechanism, what could be the impact of this impairment on the development of numerical and arithmetical cognition?

We hypothesised that due to excessive crowding, during development DD individuals could have more often perceived ensembles as being too cluttered for individual items to be clearly segregated. As a consequence, this could have led to a 'fuzzy' representation of numerosity, a less precise association between non-symbolic and symbolic numbers and ultimately less efficient arithmetical abilities.

Independent of the specific origin of visual crowding, this hypothesised chain of events should have resulted in a predictive relationship between the strength of visual crowding effects and both non-symbolic/symbolic number discrimination and calculation abilities. In particular, given that the ability to discriminate the numerosity of sparse, but not of cluttered arrays, was found to be predictive of symbolic mathematical abilities in children, we expected visual crowding to be predictive of numerosity discrimination abilities for sparse arrays. The current results did not support this prediction: we did not find a significant correlation between indices of visual crowding and weber fractions for sparse (nor for dense) dot ensembles. One reason why we might have failed to find such an effect might

be related to the fact that we tested adult participants, who might have developed compensatory strategies for dealing with numerosity tasks. For example, participants might have based their decisions on some other visual features, such as the total contrast, density or surface area which in the current experiment increased with numerosity. This might also explain why we only observed a trend for significance when comparing weber fractions for sparse arrays between the DD and the control group. This interpretation fits well with the results reported by previous studies: weber fractions measured with a non-symbolic numerosity discrimination task in DD children differed from controls only when the non-numerical dimensions varied incongruently with the numerical ones (e.g. when the numerically larger set had smaller item surface area compared to the other set) so that it could not be used as a reliable proxy for numerosity (Bugden & Ansari, 2016). Other studies found that numerosity judgments in both DD children (Szűcs, Devine, Soltesz, Nobes, & Gabriel, 2013) and adults (Castaldi, Mirassou, Dehaene, & Piazza, 2018b) were subject to enhanced congruency effects and were strongly biased by non-numerical dimensions (such as, total surface area, edge length, average item size), compared to age-matched control groups. While the current results do not allow us to affirm the link between degree of crowding and non-symbolic numerical acuity, future studies should test for such a link in children with and without DD, as well as under a wider range of controls for non-numerical quantitative properties. Ideally, such studies should also incorporate reaction time recording during the number discrimination task and verify whether there is a relation between the reaction time cost induced by crowding and the time needed to take numerical decisions on arrays of multiple items. The enhanced detrimental effect of crowding was observed in the current study when taking into account reaction times, suggesting that DD adults succeeded in compensating performance, but to do so they needed longer processing time. By measuring the weber fraction but not reaction times in the numerosity task, we might have missed one informative measure linking crowding to non-symbolic numerical abilities.

It is also important to note that the role of non-symbolic numerical abilities in the development of integer concepts has been challenged by some authors (Carey & Barner, 2019) and that the predictive relationship between non-symbolic number discrimination and formal arithmetic has sometimes not been replicated. Rather, symbolic number abilities resulted to be a more robust predictor of arithmetical skills across studies (De Smedt, Noël, Gilmore, & Ansari, 2013; Schneider et al., 2017). In the current study we observed that the orientation discrimination thresholds under crowding and the crowding index were predictive of the performance in both a digit comparison task and general math skills: participants with stronger visual crowding effects performed more poorly in digit comparison and calculation tasks. Given that in the digit comparison task, two numbers were presented far apart from each other, on the two sides of the screen, it is very unlikely that they would have been displayed within the same 'integration field'.

In the current study we ruled out the possibility that the observed across group differences in crowding effects were exclusively driven by the participants with major visuo-spatial attentional deficits, potentially identifying participants with a history of ADHD. Yet, it remains possible that some attentional weaknesses, known to be present in pure DD individuals (Askenazi & Henik, 2010), although potentially not strong enough to result significantly different with respect to the control group when measured with the visual search test used here, might have contributed to the excessive crowding effects observed. Related to this interpretation, another hypothesis that can be advanced to explain the observed relation between crowding and numerical abilities is that attentional weaknesses might have prevented the development of a sufficiently clear spatial representation of numbers. An influential hypothesis proposed that numbers are internally represented along a spatially oriented number line (Dehaene, Bossini, & Giraux, 1993; Dehaene, 2003). According to this view, spatial associations are relevant both for understanding the meaning of numerical values, which would be conveyed by their position on the number line, as well as during calculation, which is thought to involve shifts of spatial attention along the number line (Hubbard, Piazza, Pinel, & Dehaene, 2005). There is evidence that impairments of visuo-spatial attention affect this numerical space: neuropsychological studies on patients with hemi-spatial representational neglect found that they perform extremely poorly in number bisection tasks, most likely because the deficit in orienting visuo-spatial attention to the contralesional hemispace also extended to the internal representation of the number line (Zorzi, Priftis, & Umiltà, 2002, Zorzi, Priftis, Meneghello, Marenzi, & Umiltà, 2006). If the excessive crowding effects observed in the current study in the DD group reflect a limited spatial resolution of attention, then they might have hampered the development of a clear internal representation of numbers in space. Numbers might be 'too crowded' along the internal mental number line to be sharply sampled and manipulated during number comparison or calculation tasks. The possibility that crowding may degrade internal representations, and not only visual percepts, is also supported by a recent study showing that crowding also affects visual working memory contents (Tamber-Rosenau, Fintzi, & Marois, 2015). Such 'representational crowding' might set limits to the ability of the DD individuals to precisely select the two numerical values to be compared on the number line

and/or to keep them in memory during the comparison or calculation process.

It has been previously suggested by some authors that the difficulties in calculation observed in DD might be more related to impaired general executive functions such as attention, working/long-term memory and inhibition rather than to conceptual knowledge of number and arithmetic (Bull & Scerif, 2001; Geary, 2004). The relation between excessive crowding effects and numerical and calculation abilities observed in the current study could be interpreted as an evidence in support of that model of DD, with no need to evoke any domain specificity. However, arguing against such a pure domain-general hypothesis, it has been shown that training general attention in DD adults (using video-games) improved the orienting system, but did not improve the difficulties in the DD group in arithmetic or in numerical processing (Ashkenazi & Henik, 2012). Neither did specifically boosting the alerting system with brief auditory cues prior to an estimation task increase the smaller than normal subitizing range or accuracy in DD individuals (Gliksman & Henik, 2019). Moreover, domain-general deficits in DD are sometimes reported to be specific to the numerical stimuli. For example, one study found that verbal working memory deficits in DD were stronger when digits were tested with respect to letters or words (Peng & Fuchs, 2016). Others found that inhibition deficits in DD appeared only with Stroop paradigms involving digits, but not letters or geometric features (Wang, Tasi, & Yang, 2012). In ensemble discrimination tasks, enhanced interference from the task-irrelevant dimensions manifested only when DD subjects were required to judge the numerosities of the ensemble, and not when judging other quantitative features, such as its average item size (Castaldi et al., 2018b). Overall, this evidence suggests that, although DD may present some weakness in domain-general functions, these may not uniquely explain the specific numerical difficulties in DD. Rather it is more likely that weak domain-general functions might interact with an impaired number specific magnitude system. Using digit stimuli, some studies found that flankers might influence target detection not only because of their perceptual similarity, but also because of their semantic closeness: responses are facilitated when a target number is surrounded by numerically congruent flankers and asymmetries in flanker-target interference occur when magnitude or parity of the target (Huckauf, Knops, Nuerk, & Willmes, 2008), but not its physical characteristics, need to be extracted (Patro & Huckauf, 2019). To compare domain-general vs domain-specific attentional deficits in DD, future studies should test whether DD individuals present even stronger crowding effects when using numerical with respect to non-numerical stimuli (such as the Gabors used here). Moreover, in order to test our hypothesis that

crowding impacts numerical cognition by preventing the development of a clear internal number line, it would be interesting to test whether pre-cueing attention to the operands' locations along the number line would improve performance in calculation tasks by increasing the 'mental critical spacing' between numbers.

The presence of abnormal crowding in DD may seem at odds with the lack of strong difficulties in the reading and visual search tests, at least for the majority of our DD participants. This could further support the idea that the participants tested in the current experiment suffered from a form of 'representational', rather than perceptual, crowding. An alternative possibility is that crowding might differentially affect parafoveal and peripheral vision in DD. Martelli et al. (2009) reported that abnormal crowding affected all eccentricities (4–12 visual degrees) in dyslexic children, however four children presented a disproportionally large critical spacing at 4 degrees which decreased with eccentricity. Parafoveal vision is obviously the most important region for reading and visual search skills, and given that most of our DD participants were not strongly impaired in tasks measuring these abilities, it is tempting to hypothesize that crowding might affect peripheral vision more than central vision in DD individuals. Further studies are needed to verify whether there is an inverted pattern across these two neurodevelopmental disorders, with excessive parafoveal and peripheral crowding characterizing developmental dyslexia and dyscalculia respectively. Ideally, to address this and many other questions left open from this study, the impact of eccentricity on crowding should be investigated in children and correlated with the performance in non-symbolic number discrimination tasks in central versus peripheral vision.

One limitation of the current study is the sample size, with only 15 DD participants included in the study, and as many in the control group. Crowding effects were so strong for three DD participants that they were not able to perform the orientation discrimination task in the presence of flankers and therefore could not to be included in further analyses. The present results should therefore be taken with some caution and replicated in future studies with larger sample sizes. Yet, according to the Bayesian analysis performed, the different crowding measures reported in the current study were always at least positively, but also in some cases substantially, strongly and very strongly in support of excessive crowding effects in DD participants compared to the control group.

In conclusion, this study provides a first report of abnormal crowding in DD individuals, which can be found independently of pronounced associated reading and visual-attention deficits. Which are the exact mechanisms underlying the excessive crowding effects and how they contribute to the development of

numerical and arithmetical difficulties in DD needs to be explored further in future studies. Perhaps, similarly to what was previously observed in dyslexic individuals (Martelli et al., 2009), excessive crowding may contribute to create numerical difficulties in DD, although not fully accounting for it.

*Keywords: developmental dyscalculia, visual crowding, number perception, visual attention*

Commercial relationships: none.
Corresponding author: Elisa Castaldi.
Email: elisa.castaldi@gmail.com.
Address: Department of Translational Research and New technologies in Medicine and Surgery, University of Pisa, Via S. Zeno 31, 56127, Pisa, Italy.

[*]EC and MT contributed equally to this work.

## References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Publishing.

Anobile, G., Cicchini, G. M., & Burr, D. (2013). Separate mechanisms for perception of numerosity and density. *Psychological Science, 25*(1), 265–270.

Anobile, G., Turi, M., Cicchini, G. M., & Burr, D. (2015). Mechanisms for perception of numerosity or texture-density are governed by crowding-like effects. *Journal of Vision, 15*, 4.

Anobile, G., Castaldi, E., Turi, M., Tinelli, F., & Burr, D. (2016a). Numerosity but not texture-density discrimination correlates with math ability in children. *Developmental Psychology, 52*, 1206–1216.

Anobile, G., Cicchini, G. M., & Burr, D. (2016b). Number As a Primary Perceptual Attribute: A Review. *Perception, 45*, 5–31.

Anobile, G., Arrighi, R., Castaldi, E., Grassi, E., Pedonese, L., M. Moscoso, P. A., . . . Burr, D. (2018). Spatial but Not Temporal Numerosity Thresholds Correlate With Formal Math Skills in Children. *Developmental Psychology, 54*, 458–473.

Ashkenazi, S., & Henik, A. (2012). Does attentional training improve numerical processing in developmental dyscalculia? *Neuropsychology, 26*, 45–56.

Askenazi, S., & Henik, A. (2010). Attentional networks in developmental dyscalculia. *Behavioral and Brain Functions, 6*, 2.

Balas, B. (2016). Seeing number using texture: How summary statistics account for reductions in perceived numerosity in the visual periphery. *Atten Percept Psychophys, 78*, 2313–2319.

Bayard, S., Erkes, J., & Moroni, C. (2009). *Test du Stroop Victoria-Adaptation Francophone*. CPCN-LR, Gignac.

van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., van Kesteren, E., Derks, K., Dablander, F., Gronau, Q. F., Kucharsk, S., . . . Komarlu Narendra Gupta, A. R. et al. (2019). A Tutorial on Conducting and Interpreting a Bayesian ANOVA in JASP. *PsyArXiv*, doi:10.31234/osf.io/spreb.

Bondarko, V. M., & Semenov, L. A. (2005). Visual Acuity and the Crowding Effect in 8- to 17-Year-Old Schoolchildren. *Hum Physiol, 31*, 532–538.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*, 177–178.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436, https://doi.org/10.1163/156856897X00357.

Brannon, E. M., Abbott, S., & Lutz, D. J. (2004). Number bias for the discrimination of large visual sets in infancy. *Cognition, 93*, B59–B68.

Bugden, S., & Ansari, D. (2016). Probing the nature of deficits in the 'Approximate Number System' in children with persistent Developmental Dyscalculia. *Developmental Science, 19*, 817–833.

Bull, R., & Scerif, G. (2001). Executive Functioning as a Predictor of Children's Mathematics Ability: Inhibition, Switching, and Working Memory. *Developmental Neuropsychology, 19*, 273–293.

Burr, D., & Ross, J. (2008). A Visual Sense of Number. *Current Biology, 18*, 425–428.

Burr, D., Anobile, G., & Arrighi, R. (2017). Psychophysical evidence for the number sense. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*, 20170045.

Butterworth, B. (2005). Developmental dyscalculia. In *Handbook of Mathematical Cognition* (pp. 455–467). Psychology Press, Hove.

Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Sciences, 14*, 534–541.

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.

Carey, S., & Barner, D. (2019). Ontogenetic Origins of Human Integer Representations. *Trends in Cognitive Sciences, 23*, 823–835.

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.

Castaldi, E., Tinelli, F., Cicchini, G. M., & Morrone, M. C. (2018a). Supramodal agnosia for oblique mirror orientation in patients with periventricular leukomalacia. *Cortex, 103*, 179–198.

Castaldi, E., Mirassou, A., Dehaene, S., Piazza, M., & Eger, E. (2018b). Asymmetrical interference between number and item size perception provides evidence for a domain specific impairment in dyscalculia. *PLOS ONE, 13*, e0209256.

Chakravarthi, R., & Bertamini, M. (2020). Clustering leads to underestimation of numerosity, but crowding is not the cause. *Cognition, 198*, 104195.

Chaney, W., Fischer, J., & Whitney, D. (2014). The hierarchical sparse selection model of visual crowding. *Frontiers in Integrative Neuroscience, 8*, 73, doi:10.3389/fnint.2014.00073.

Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica, 148*, 163–172.

Collins, E., Park, J., & Behrmann, M. (2017). Numerosity representation is encoded in human subcortex. *Proceedings of the National Academy of Sciences, 114*, E2806–E2815.

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences, 108*, 19552–19557.

De Smedt, B., Noël, M. P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education, 2*, 48–55.

Dehaene, S. (1997). *The number sense: how the mind creates mathematics*. New York: Oxford University Press.

Dehaene, S. (2003). The neural basis of the Weber – Fechner law: a logarithmic mental number line. *Trends in Cognitive Sciences, 7*(4), 145–147, doi:10.1016/s1364-6613(03)00055-x.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*, 371–396.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*, 487–506.

van Doorn, J., van den Bergh, D., Bohm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., . . . Hinne, M. et al. (2019). The JASP Guidelines for Conducting and Reporting a Bayesian Analysis (PsyArXiv), doi:10.31234/osf.io/yqxfr.

Faulkenberry, T. J., Ly, A., & Wagenmakers, E.-J. (2020). Bayesian inference in numerical cognition: A tutorial using JASP (PsyArXiv), doi:10.31234/osf.io/vg9pw.

Fias, W. (2015). Neurocognitive Components of Mathematical Skills and Dyscalculia. In D. Berch, D. Geary, & K. Mann Koepke (Eds.), *Development of Mathematical Cognition* (pp. 195–217). Amsterdam: Elsevier.

Fias, W., Menon, V., & Szucs, D. (2013). Multiple components of developmental dyscalculia. *Trends in Neuroscience and Education, 2*, 43–47.

Flom, M. C., Heath, G. G., & Takahashi, E. (1963). Contour interaction and visual resolution: Contralateral effect. *Science, 142*(3594), 979–980, doi:10.1126/science.142.3594.979.

Freyberg, J., Robertson, C. E., & Baron-Cohen, S. (2016). Typical magnitude and spatial extent of crowding in autism. *Journal of Vision, 16*, 17.

Geary, D. C. (2004). Mathematics and Learning Disabilities. *Journal of Learning Disabilities, 37*, 4–15.

Gliksman, Y., & Henik, A. (2019). Enumeration and Alertness in Developmental Dyscalculia. *Journal of Cognition, 2*, 5.

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*, 8–8.

Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2010). Crowding Changes Appearance. *Current Biology, 20*, 496–501.

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*, 665–668.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383(6598)*, 334–337.

de Hevia, M. D. (2016). Core mathematical abilities in infants. In *Progress in Brain Research* (pp. 53–74). Amsterdam: Elsevier.

de Hevia, M. D., Castaldi, E., Streri, A., Eger, E., & Izard, V. (2017). Perceiving numerosity from birth. *Behavioral and Brain Sciences, 40*, 21–22.

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience, 6*, 435–448.

Huckauf, A., Knops, A., Nuerk, H.-C., & Willmes, K. (2008). Semantic processing of crowded stimuli? *Psychological Research, 72*, 648–656.

Intriligator, J., & Cavanagh, P. (2001). The Spatial Resolution of Visual Attention. *Cognitive Psychology, 43*, 171–216.

Iuculano, T. (2016). Neurocognitive accounts of developmental dyscalculia and its remediation. In *Progress in Brain Research* (pp. 305–333). Amsterdam: Elsevier.

Iuculano, T., Tang, J., Hall, C. W. B., & Butterworth, B. (2008). Core information processing deficits in developmental dyscalculia and low numeracy. *Developmental Science, 11*, 669–680.

Izard, V., Sann, C., Spelke, E. S., Streri, A., & Gallistel, C. R. (2009). Newborn Infants Perceive Abstract Numbers. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 10382–10385.

JASP Team. (2020). JASP (Version 0.13.1) [Computer software], https://jasp-stats.org/.

Kwon, M., Legge, G. E., & Dubbels, B. R. (2007). Developmental changes in the visual span for reading. *Vision Research, 47*(22), 2889–2900, doi:10.1016/j.visres.2007.08.002.

Lefavrais, P. (1967). Test de l'Alouette. Editions du Centre de Psychologie Appliquée, Paris.

Legge, G. E., Mansfield, J. S., & Chung, S. T. L. (2001). Psychophysics of reading XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Research, 41*(6), 725–743, doi:10.1016/s0042-6989(00)00295-9.

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research, 48*, 635–654.

Libertus, M. E., Starr, A., & Brannon, E. M. (2014). Number trumps area for 7-month-old infants. *Developmental Psychology, 50*, 108–112.

Manassi, M., & Whitney, D. (2018). Multi-level Crowding and the Paradox of Object Recognition in Clutter. *Current Biology, 28*, R127–R133.

Martelli, M., Di Filippo, G., Spinelli, D., & Zoccolotti, P. (2009). Crowding, reading, and developmental dyslexia. *Journal of Vision, 9*, 14–14.

Mejias, S., Grégoire, J., & Noël, M.-P. (2012). Numerical estimation in adults with and without developmental dyscalculia. *Learning and Individual Differences, 22*, 164–170.

Moll, K., & Jones, M. (2013). Naming Fluency in Dyslexic and Nondyslexic Readers: Differential Effects of Visual Crowding in Foveal, Parafoveal, and Peripheral Vision. *Quarterly Journal of Experimental Psychology, 66*, 2085–2091.

Montani, V., Facoetti, A., & Zorzi, M. (2015). The effect of decreased interletter spacing on orthographic processing. *Psychon Bull Rev, 22*, 824–832.

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., Raphael, S., Tibber, M. S., & Dakin, S. C. (2014). A texture-processing model of the "visual sense of number." *Proceedings of the Royal Society B: Biological Sciences, 281*, 20141137–20141137.

Noël, M. P., & Grégoire, J. (2015). *TEDI-MATH Grands: test diagnostique des compétences de base en mathématiques du CE2 à la 5e*. Pearson France ECPA: Montreuil, France. 145, http://hdl.handle.net/2078.1/167790.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4*, 739–744.

Patro, K., & Huckauf, A. (2019). Asymmetries in flanker-target interference at different levels of number processing. *Acta Psychologica, 201*, 102938.

Pelli, D. G. (2008). Crowding: a cortical constraint on object recognition. *Current Opinion in Neurobiology, 18*, 445–451.

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience, 11*, 1129–1135.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4*, 12.

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision, 7*, 20.

Peng, P., & Fuchs, D. (2016). A meta-analysis of working memory deficits in children with learning difficulties: Is there a difference between verbal domain and numerical domain? *Journal of Learning Disabilities, 49*, 3–20.

Perea, M., Panadero, V., Moret-Tatay, C., & Gómez, P. (2012). The effects of inter-letter spacing in visual-word recognition: Evidence with young normal readers and developmental dyslexics. *Learning and Instruction, 22*, 420–430.

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences, 14*, 542–551.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., Dehaene, S., . . . Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition, 116*, 33–41.

Piazza, M., Pica, P., Izard, V., Spelke, E., & Dehaene, S. (2013). Education Enhances the Acuity of the Nonverbal Approximate Number System. *Psychological Science, 24*, 1037–1043.

Ross, J. (2010). Vision senses number directly. *Journal of Vision, 10*, 1–8.

Rousselle, L., & Noël, M. P. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number magnitude processing. *Cognition, 102*, 361–395.

Rubinsten, O. (2009). Co-occurrence of developmental disorders: The case of Developmental Dyscalculia. *Cognitive Development, 24*, 362–370.

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., . . . De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Dev Sci.*, *20*, e12372.

Spinelli, D., De Luca, M., Judica, A., & Zoccolotti, P. (2002). Crowding Effects on Word Identification in Developmental Dyslexia. *Cortex, 38*, 179–200.

Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision, 5*, 1024–1037.

Szűcs, D. (2016). Subtypes and comorbidity in mathematical learning disabilities. In *Progress in Brain Research* (pp. 277–304). Amsterdam: Elsevier.

Szűcs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex, 49*, 2674–2688.

Tamber-Rosenau, B. J., Fintzi, A. R., & Marois, R. (2015). Crowding in Visual Working Memory Reveals Its Spatial Resolution and the Nature of Its Representations. *Psychological Science, 26*, 1511–1521.

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research, 32*, 1349–1357.

Valsecchi, M., Toscani, M., & Gegenfurtner, K. R. (2013). Perceived numerosity is reduced in peripheral vision. *Journal of Vision, 13*, 7–7.

Wang, L.-C., Tasi, H.-J., & Yang, H.-M. (2012). Cognitive inhibition in students with and without dyslexia and dyscalculia. *Research in Developmental Disabilities, 33*, 1453–1461.

Watson, A. B., & Pelli, D. G. (1983). Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*, 113–120.

Whitney, D., & Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences, 15*, 160–168.

Wilson, A. J., Andrewes, S. G., Struthers, H., Rowe, V. M., Bogdanovic, R., & Waldie, K. E. (2015). Dyscalculia and dyslexia in adults: Cognitive bases of comorbidity. *Learning and Individual Differences, 37*, 118–132.

Yeshurun, Y., & Rashal, E. (2010). Precueing attention to the target location diminishes crowding and reduces the critical distance. *Journal of Vision, 10*, 16–16.

Yu, D., Cheung, S.-H., Legge, G. E., & Chung, S. T. L. (2007). Effect of letter spacing on visual span and reading speed. *Journal of Vision, 7*, 2.

Zorzi, M., Priftis, K., & Umiltà, C. (2002). Neglect disrupts the mental number line. *Nature, 417*, 138–139.

Zorzi, M., Priftis, K., Meneghello, F., Marenzi, R., & Umiltà, C. (2006). The spatial representation of numerical and non-numerical sequences: Evidence from neglect. *Neuropsychologia, 44*, 1061–1067.

Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M., Bravar, L., George, F., Pech-Georgel, C., . . . Ziegler, J. C. (2012). Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences, 109*, 11455–11459.