Spontaneous representation of numerosity in typical and dyscalculic development

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Special issue: Research report

Spontaneous representation of numerosity in typical and dyscalculic development

Guido Marco Cicchini [a,*], *Giovanni Anobile* [b] *and David C. Burr* [c,d]

[a] *Institute of Neuroscience, National Research Council, Pisa, Italy*
[b] *Department of Developmental Neuroscience, IRCCS Stella Maris Foundation, Pisa, Italy*
[c] *Department of Neuroscience, Psychology, Pharmacology and Child Health, University of Florence, Florence, Italy*
[d] *School of Psychology, University of Sydney, Sydney, Australia*

ARTICLE INFO

Article history:
Received 19 December 2017
Reviewed 7 April 2018
Revised 8 September 2018
Accepted 14 November 2018
Published online 28 December 2018

Keywords:
Approximate number system
Developmental dyscalculia
Numerosity perception
Number sense
Numerical cognition

ABSTRACT

Animals including humans are endowed with a remarkable capacity to estimate rapidly the number of items in a scene. Some have questioned whether this ability reflects a genuine sense of number, or whether numerosity is derived indirectly from other co-varying attributes, such as density and area. In previous work we have demonstrated that adult observers are more sensitive to changes in numerosity than to area or density, particularly changes that leave numerosity constant, pointing to a spontaneous sensitivity to numerosity, not attributable to area and density. Here we extend this line of research with a novel technique where participants reproduce the size and density of a dot-array. They were given no explicit instructions of what to match, but could regulate freely all combinations of area and density by trackpad. If the task is mediated by matching separately area and texture-density, the errors in the two attributes have to be independent. Contrarily to this prediction, we found that errors in area and density were negatively correlated, suggesting that subjects matched numerosity, rather than area and density. We employed this technique to investigate processing of number in adolescents with typical and low math abilities (dyscalculia). Interestingly, we found that dyscalculics also reproduced numerosity rather than area or density. However, compared to typicals, dyscalculics had longer reaction times, a tendency to rely also on area, and their performance did not improve over sessions. Taken together, the data demonstrate that numerosity emerges as the most spontaneous and sensitive dimension, supporting the existence of a dedicated number sense and confirm numerosity atypicalities in dyscalculia.

* *Corresponding author.*
E-mail address: cicchini@in.cnr.it (G.M. Cicchini).

https://doi.org/10.1016/j.cortex.2018.11.019

1. Introduction

When serial counting is prevented, humans and many other animals can estimate roughly but rapidly the number of visual objects (Butterworth, 1999; Dehaene, 1992, 2011) by a process that has been termed the "Approximate Number System (ANS)". The importance of such a system is obvious: from sensing the number of coins in a jar, fruits on a tree or enemies on the field, it is clear that the capacity to extract number from the environment is important for crucial decisions for humans in both modern and ancient societies, and throughout the animal kingdom. Indeed, there is evidence that the system is functional 2 days after birth (Izard, Sann, Spelke, & Streri, 2009).

Between 5 and 8% of school age children suffer from dyscalculia, a neurodevelopmental disorder characterized by an impairment of learning mathematical concepts. Interestingly this deficit is often accompanied by severe difficulties in estimating and comparing numerosity (Piazza, Facoetti, Trussardi, Berteletti, Conte, & Lucangeli, 2010; Mazzocco, Feigenson, & Halberda, 2011; Pinheiro-Chagas et al., 2014) leading some authors to hypothesize that an intact number sense is a core prerequisite for later mathematical acquisition (Butterworth, 1999; Piazza, 2010; Butterworth, Varma, & Laurillard, 2011; Dehaene, 2011).

But how visual numerosity is perceived is a highly debated issue. The dispute arises from the inevitable fact that in nature, numerosity correlates with many other stimulus properties such as density, surface area, luminance and so on. All these features carry congruent information and can thus be used as a proxy for numerosity judgements (Dakin, Tibber, Greenwood, Kingdom, & Morgan, 2011; Gebuis, Cohen Kadosh, & Gevers, 2016; Gebuis & Reynvoet, 2012; Leibovich, Katzin, Harel, & Henik, 2016).

Numerosity sensitivity is usually ascertained by asking people to explicitly compare two briefly presented sets, and determining the minimal detectable numerical difference (discrimination threshold). These measurements show that numerosity, like many other visual attributes, follows Weber's Law (discrimination thresholds scale linearly with numerosity), at least over a limited range (Anobile, Cicchini, & Burr, 2014; Dehaene, 2011; Jevons, 1871; Ross, 2003). Discrimination thresholds decrease during development and show a developmental delay in children suffering from developmental dyscalculia (Piazza et al., 2010).

Numerosity is also susceptible to sensory adaptation, another ubiquitous feature of visual perception: after observing for a few seconds a large number of dots, a subsequent set of dots of moderate numerosity will appear to contain fewer elements than are there: and vice versa for low numerosity adapters (Aagten-Murphy & Burr, 2016; Burr, Anobile, & Turi, 2011; Burr & Ross, 2008). Given these properties, it has been suggested that numerosity can be considered a primary visual feature, like colour or motion (Burr & Ross, 2008). However, this idea has been challenged by several authors who suggest that numerosity may be not perceived directly but recomputed from other correlated visual proprieties, such as texture-density and stimulus area (Dakin et al., 2011; Durgin, 2008). For example, items displayed within a larger patch area appear more numerous, suggesting a shared metric. More recently Leibovich et al. (2016) have advanced a more extreme hypothesis, questioning the existence of a specific number sense, suggesting instead a more generalized sense of magnitude encoding many continuous features including area and density. This idea is in line with many other suggestions that numerosity is not sensed directly, but indirectly via texture-density mechanisms (Dakin et al., 2011; Durgin, 2008; Tibber, Greenwood, & Dakin, 2012).

The past decade has seen a lively debate on whether numerosity is sensed directly, by what could be called a *number sense* (Ross & Burr, 2010; Stoianov & Zorzi, 2012; Harvey, Klein, Petridou, & Dumoulin, 2013; Anobile, Castaldi, Turi, Tinelli, & Burr, 2016; Park, DeWind, Woldorff, & Brannon, 2016; Anobile, Cicchini, Pomè, & Burr, 2017; Ferrigno, Jara-Ettinger, Piantadosi, & Cantlon, 2017; Odic, 2017), or indirectly (Durgin, 2008; Dakin et al., 2011; Tibber et al., 2012, 2013; Morgan, Raphael, Tibber, & Dakin, 2014; Leibovich et al., 2016). One possible resolution has been to suggest that both texture-density and numerosity mechanisms may operate, depending on the density of the displays (Anobile et al., 2014). This would imply that there exist three different but overlapping regimes of number perception. For very low numbers, less than four, attention-dependent *subitizing* mechanisms operate, rapidly and errorlessly (Kaufman & Lord, 1949; Egeth, Leonard, & Palomares, 2008; Olivers & Watson, 2008; Railo, Koivisto, Revonsuo, & Hannula, 2008; Vetter, Butterworth, & Bahrami, 2008; Burr, Turi, & Anobile, 2010; Anobile, Turi, Cicchini, & Burr, 2012). For items greater than four, when there are too many to be appraised immediately, but they are sparse enough to be perceptually segregable (Anobile, Cicchini, & Burr, 2016), the approximate number system prevails (Feigenson, Dehaene, & Spelke, 2004). Performance in this regime has been shown to obey Weber's law (Anobile et al., 2014), and seems to be quite robust to manipulation of low-level visual quantities, such as the area of the dot patch (see Fig. 8 in Cicchini, Anobile, & Burr, 2016). For denser item sets, too dense for individual items to be discerned, texture-density mechanisms come into play, where geometrical properties of the display such as its size and its local density tend to dominate discrimination. This regime is characterized by a square-root relationship rather than Weber's law (Anobile et al., 2014). The transition between the estimation and texture-density regimes depends on eccentricity, typically about .5 items per square degree in central viewing, decreasing with increasing eccentricity (Anobile, Turi, Cicchini, & Burr, 2015).

There is now a good deal of evidence that for reasonably sparse items, numerosity is sensed directly (Anobile, Cicchini, et al., 2016; Burr, Anobile, & Arrighi, 2017). One line of studies demonstrates the existence of a numbersense for temporal sequences, where texture mechanisms could not possibly operate (Arrighi, Togoli, & Burr, 2014; Anobile, Arrighi, Togoli, & Burr, 2016). The sequences can be in any sensory modality, visual, auditory or haptic. More interestingly, there are clear interactions between temporal sequences and spatial arrays of objects, pointing to a generalized number sense transcending both sensory modality, and also space and time.

Another approach has been to employ an assumption-free technique that tests directly which is the more spontaneous

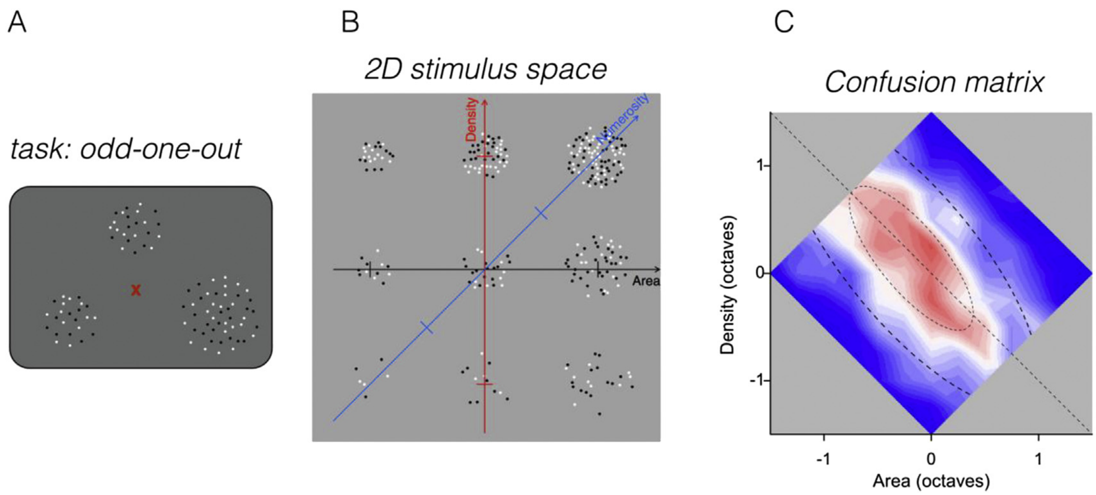

Fig. 1 — Spontaneous discrimination of numerosity. A) Illustration of the stimuli used in the odd-one-out task of Cicchini et al. (2016). Two dot-clouds were identical, defined by the origin of the space (24 dots, 40 deg^2). The target (lower right) was taken from many different sample points in the space (in this case, 48 dots, 80 deg^2). B) Stimulus space, indicating changes in area (horizontal axis), density (vertical axis) and this number (diagonal). Ticks indicate 1 octave changes. C) Two dimensional psychometric functions confusion matrices on a space spanned by log-area and log-density, for sparse standard stimuli (24 dots, 40 deg^2, .6 dots/deg^2). The heat map refers to the interpolated percent correct for discriminating various points sampled over the coloured area from the origin (with 33% chance guessing rate). The ellipses are fitted to pass through 50% and 75% correct judgement. Modified with permission from Cicchini et al. (2016).

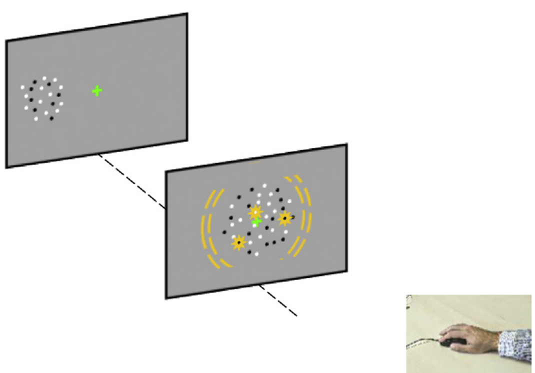

Fig. 2 — Dot-cloud reproduction. After a brief presentation of a sample dot-cloud in the left hemifield, subjects were prompted with a new dot-cloud in the centre of the screen, which they could edit with mouse or track-pad movements. Horizontal movements increased or decreased the area (logarithmically), exposing new or removing old dots, leaving density unchanged. Vertical movements increased or decreased density, adding or subtracting dots from the pattern without changing area. Once satisfied with the match, subjects clicked the mouse and proceeded to the next trial. See on-line movie 1.

and sensitive dimension, numerosity or density and area (Cicchini et al., 2016). Fig. 1 illustrates the technique and results. Subjects were presented with three dot-clouds and asked to identify which of them was different (in any respect). Stimuli were defined within a two-dimensional space, spanned by log-area (abscissa) and log-density (ordinate), where log-numerosity follows the positive diagonal. Two dot-clouds

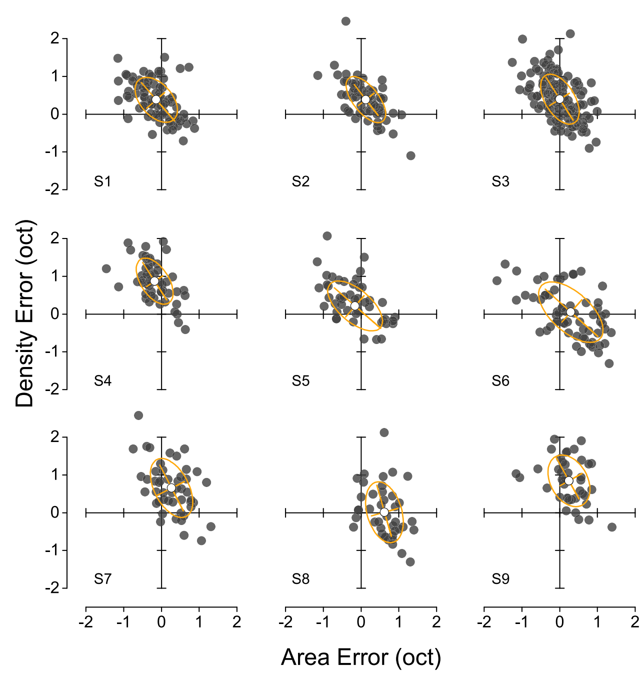

Fig. 3 — Reproduction errors for a low-density cloud of dots. Errors in reproduction on a logarithmic plot of density against area. The origin refers to the reference (area 78.5 deg^2, 12 dots), and each dot is an individual trial of a subject, plotted as the ratio between actual and reproduced area and density, in octaves. Each panel shows data of a single subject obtained over the course of a minimum 3 sessions. For each subject we overlay the ellipse encircling the full width half maximum of the 2D gaussian fit to the data.

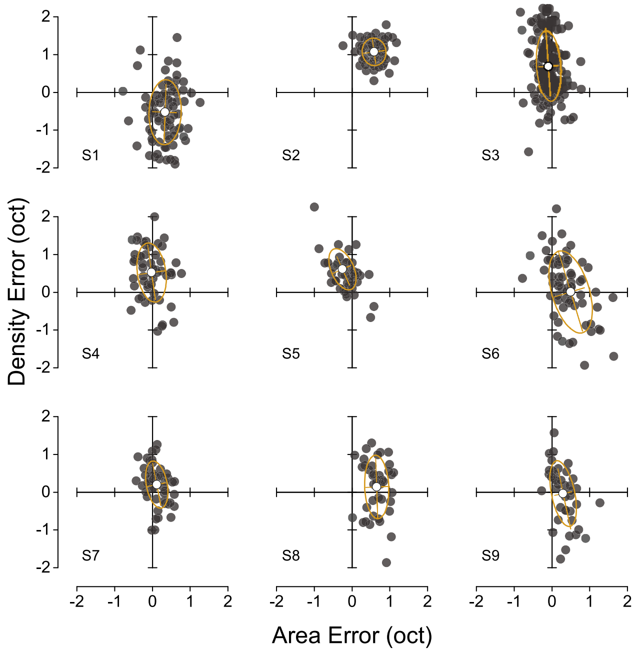

Fig. 4 – Reproduction errors for a high-density cloud of dots. Same as Fig. 3, but for high density stimuli (reference area 78.5 deg^2, reference numerosity 128 dots).

were identical, defined by the origin of the space, the third cloud differing in area and/or density (hence number). Subjects were very sensitive to changes that resulted in increased or decreased numerosity, but particularly insensitive to changes that left numerosity constant, such as reciprocal increase and decrease in area and density. The resultant two-dimensional psychometric functions followed an ellipsoidal shape, falling along the line of constant numerosity, strong evidence that perceived numerosity with sparse items cannot be derived from density, as thresholds for both density and area are far worse than those for number.

While the task used for the data of Fig. 1 was very revealing about the mechanisms underlying numerosity perception, it requires a great deal of data to determine the optimal axis for sensitivity, making it difficult to use with children and clinical populations. In the current study we expand on this work by introducing a simplified assumption-free technique where observers are required to actively reproduce a cloud of dots on each trial. We tested this task with typically developing adults, as well as with two cohorts of adolescents, with typical or impaired math abilities (dyscalculia). Similarly to our previous research, the new assumption-free task showed that participants spontaneously reproduce item numerosity with greater precision than other features such as patch area and density. This task yields further confirmation of the existence of the number sense, and provides a new, fast, simple and engaging technique to test numerosity perception in both typical and atypical development, and reinforces the idea that numerosity is a primary visual attribute that emerges spontaneously when analysing the visual scene.

2. Methods

2.1. *Participants*

We recruited two groups of healthy volunteers, one comprising adult observers (N = 9, 4 males, ranging from 26 to 41 years old, average 30.2), one comprising younger participants (32 with typical math abilities, 18 with diagnosis of dyscalculia – see Table 1 for descriptive statistics). We also tested three very young children, 5-6 years old. All participants had normal or corrected-to-normal vision. Adolescent typicals were recruited in secondary schools and were examined for any developmental disorder or special needs at school. The adolescent dyscalculics were selected from a sample referred to a local clinical service (Stella Maris Scientific Institute, Pisa) for learning disabilities. Diagnoses were made by an interdisciplinary team, following DSM-V diagnostic criteria; reading and math skills were measured using age-standardized Italian batteries. Summary description of individual tests is provided in Table 1.

Experimental procedures were approved by the local ethics committee (*Comitato Etico Pediatrico Regionale—Azienda Ospedaliero-Universitaria Meyer*—Florence, Italy; protocol n. GR-2013-02358262) and are in line with the declaration of Helsinki. All subjects gave written informed consent.

2.2. *Stimuli and procedure*

Subjects sat 57 cm from a calibrated LCD monitor (Acer, 24" 1920 × 1080, 60 Hz) spanning 52° × 29° of visual field in a dimly lit quiet room. Stimuli were generated via Matlab (Mathworks, Natwick, MA) and the graphic visualization routines Psychtoolbox (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997).

Stimuli were random clouds of dots (.2° diameter, 50% white, 50% black in order to balance luminance) defined by a specific numerosity, density and area. Reference dot clouds were presented 12° left of fixation for 500 msec while subjects maintained central fixation. After a 1 sec pause subjects were presented with a sample cloud of dots and asked to match the area and density to the sample they had seen, using a mouse or trackpad (depending on subject preference) (see Fig. 2 and online movie 1). Horizontal mouse movements varied patch area (in a logarithmic manner), vertical mouse movements log density. A diagonal movement along the +45° axis changed both density and area, hence numerosity; a movement along the orthogonal axis changed density and area while keeping numerosity constant. Participants were told about area and density, but numerosity was not explicitly mentioned. Adult observers were told that they would see a "cloud of dots", and their task was "to reproduce it to be as similar as possible to the image you first saw; to do this, move the mouse (or finger on the trackpad) and edit this image (the reference); moving left/right you make it larger or smaller; going up and down you fill or empty it". Subjects were allowed to choose which interface (mouse or trackpad) they felt most comfortable with. When collecting data with adolescents, subjects were simply told that they would be shown an "image", without any

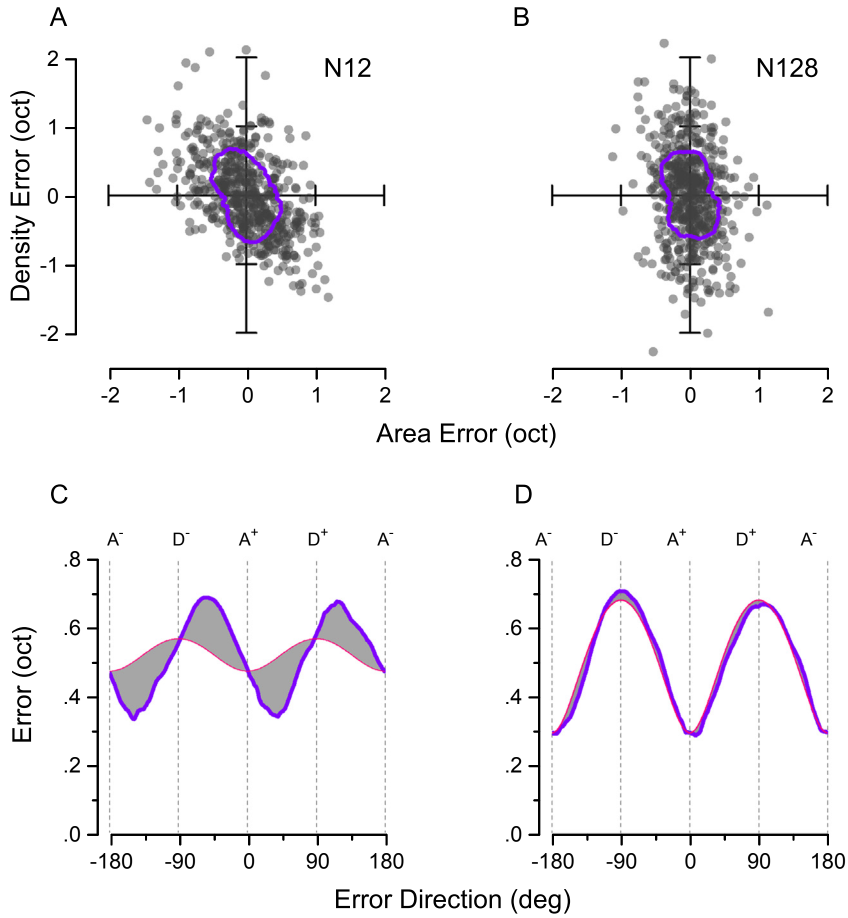

Fig. 5 — Reproduction errors at low and high density. A & B) Errors in the reproduction task for low density and high density dot-clouds (numerosities 12 and 128), for the aggregate observer (all nine typical adults, each aligned to the origin by subtracting the average error). Single dots indicate deviations in a single trial, the purple contour indicates the 84th percentile along each direction. C & D) Reproduction errors along each direction for low and high density clouds. Purple line indicates actual error, replotted from the polar plot of figs A & B in Cartesian coordinates. Thin pink lines indicate predictions for an observer who derives number as the product of area and density, calculated by propagation of errors.

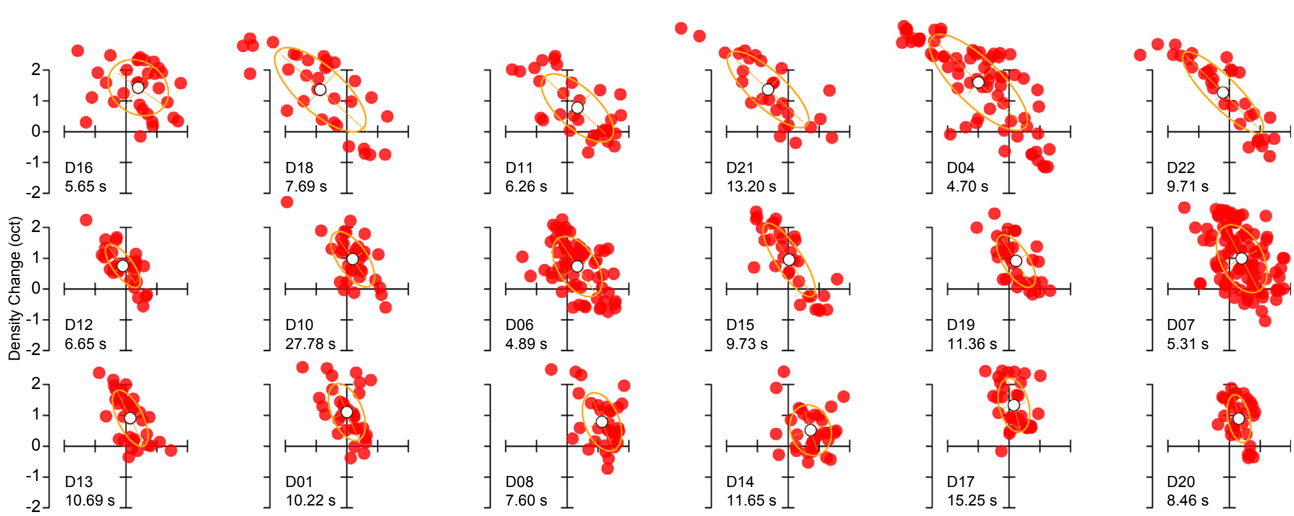

Fig. 6 — Performance of 18 dyscalculic adolescents in reproduction a low numerosity patch. Using the same conventions of above we plot here error distributions for 18 dyscalculics reproduction of dot cloud of 12 dots. Subjects are sorted according to the distribution tilt, average reaction times are reported in the inset.

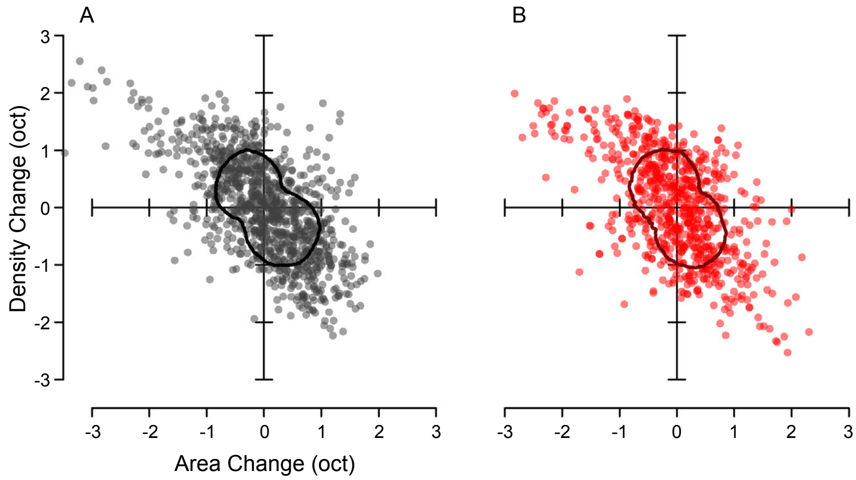

Fig. 7 — Aggregate observers for the two adolescent groups at low density. A–B Following the same convention of Fig. 5 we plot errors in the reproduction task for a low density (12 dots), for the aggregate observer obtained removing the average error of each subject. A) Typical pre-adolescents B) Dyscalculic pre-adolescents.

mentioning of the dots that made it up. Before data acquisition subjects were invited to familiarize with the way they could alter the patch characteristics and after 15 training trials data collection begun. Adult participants were expert psychophysical observers and were invited for a brief data collection. Each contributed with at least 3 sessions of 30 trials with some contributing up to 6 sessions.

Supplementary video related to this article can be found at https://doi.org/10.1016/j.cortex.2018.11.019.

In order to obtain a smooth transition between the possible clouds chosen by the subject, patterns were pre-generated at the beginning of each trial and dots were added to or removed at each mouse movement from the patch present on the screen. Mouse movements of 1.25 cm brought about changes of one octave (i.e., a factor of 2) in either area or density. The initial configuration of the matching dot-cloud was a fresh random pattern, differing from the reference stimulus by a random amount between ±2 octaves in both area and density.

Reference dot clouds were generated to cover a circle of approximately 5° radius. In order to equalize across presentations, the convex hull was calculated and the patch rescaled so that its area was equal to 78.5 deg^2. The reference stimuli contained either 12 or 128 dots (intermingled in each session) for the adult and the pre-schooler group and 12 and 24 for the adolescents. Densities ranged from .15 to 1.6 items/deg^2. Subjects were allowed a maximum of 10 sec for adjustment, and were warned if they exceeded the time limit. In any event, all trials were included in the final analysis, even if past the time limit.

Code for generating stimuli along with instructions for data analysis are made available on the authors website (http://win.pisavisionlab.org/downloads/NumberReproduction.zip).

2.3. Analysis

The area and density of the reproduced patterns of each trial were plotted on a two-dimensional logarithmic space with the abscissa representing the ratio of the match and reference area, and the ordinate the ratio of match and reference density (see examples in Fig. 3). For all subjects this was done separately for the two base numerosities.

The data were analysed in two ways. First we asked whether the area and density were independent or correlated, by calculating the covariance matrix between the two dimensions. We then extracted the eigenvalues and the eigenvectors of the covariance matrices which correspond to the principal components of the data. From these we calculated the best-fitting two-dimensional Gaussian. This is plotted in the figures as the ellipse passing though half the maximum of the probability density function.

In a second complementary analysis we calculated response variability along all polar angles. We projected the dot-cloud around 360 degrees of angles (centred at the mean of the cloud), and calculated the 84th percentile of the responses in that direction.

To extract meaningful summary statistics of error we also calculated two measures of dispersion. One is the average scatter, which is the root-mean-square of the distance between each reproduced item and the average reproduction. This measure is an extension of the calculation of "standard deviation" allowing for the fact that responses are on a bidimensional space and that the overall change is the sum of the changes along the two dimensions (not the Pythagorean sum). A second measure of error is the Average Error which is the root-mean-square of the distances of responses from the actual stimulus. Again distances have been expressed as change (i.e., the sum of the absolute change in area and density between the two visual arrays).

2.4. Forced-choice discrimination

We also measured number discrimination with a traditional 2AFC paradigm. Two dot-clouds (.25° diameter, half-white, half-black) enclosed in a 5° radius, were presented at 12° eccentricity to the left and right of fixation for 250 msec. The reference stimulus comprised 24 dots and the probe varied from 12 to 48 dots, following an adaptive routine (Watson & Pelli, 1983). Subjects indicated which of the two was more

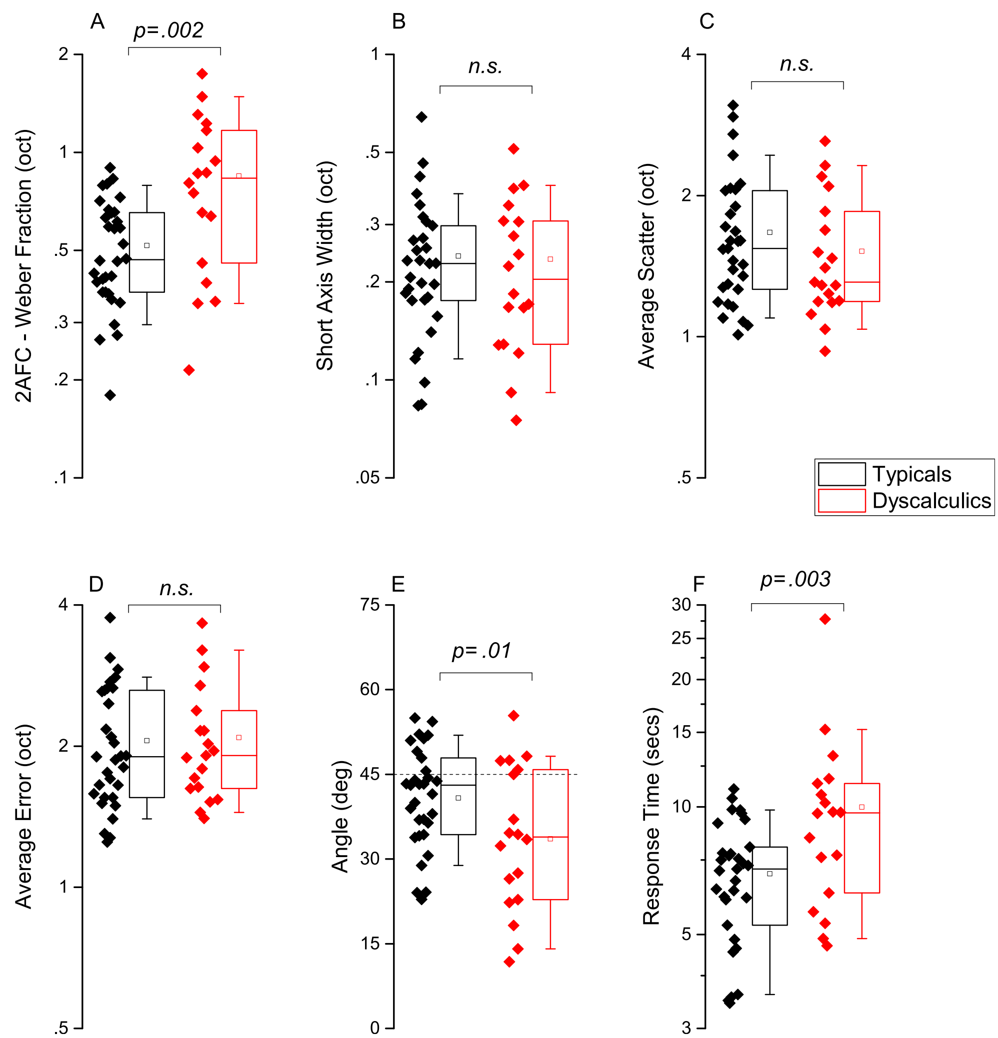

Fig. 8 – Summary statistics for the typical and dyscalculic adolescents For each individual we plot Weber Fractions from explicit numerosity comparison (A) and 5 parameters extracted from the error clouds of the reproduction task (Short Axis Width (B), Average Scatter (C), Average Error (D), Main Angle of the ellipse (E) along with the numerosity axis (dashed line) and Response Time (F)). Black are typicals, red and dyscalculics. Isolated dots indicate a single subject, box-whisker contains quartiles, means, 10th and 90th percentile.

numerous, completing two sessions of 45 trials each. Numerosity discrimination sessions were run after the array reproduction sessions to avoid that the task instructions of the forced choice ("which is more numerous") cued subjects to pay particular attention to numerosity in the reproduction trials. Data were fitted with cumulative gaussian psychometric curves extracting the Point of Subjective Equality (PSE: the median of the curve) and the Just Noticeable Difference (JND: the numerosity necessary to go from 50% to 84% 'more numerous' responses). JND divided by PSE gives the Weber Fraction, an index of sensitivity in the task.

3. Results

3.1. *Typical adult observers*

Nine typical adult observers were asked to reproduce the area and density of random dot patterns containing either 12 or 128 dots. Fig. 3 plots the errors for the low density, 12-dot patterns. The origin of the plots refers to the physical area and density of the reference stimulus, and each point to the signed logarithmic error in area (abscissa) and density (ordinate) of each individual response. As is clear from inspection, for most subjects the response distribution tends to lie along the negative diagonal, the axis of constant numerosity and the short axis of the fitted ellipses lie near the numerosity axis (mean = +36°, SD = 11°). This suggests that the lowest errors were for variations in numerosity, and the largest when numerosity was constant. Analogous to our previous work with odd-one-out discriminations, these ellipses can be considered zones of confusion, within which stimuli are indistinguishable. For most subjects the most distinguishable stimuli were those that varied in numerosity.

To quantify the results, we extracted for each subject summary statistics that describe the slant and the inherent structure of the data. We first calculated the covariance between area and density reproduction errors. If subjects

Table 1 – Descriptive statistics of dyscalculics and control adolescent.

Group	Age (std) [range]	WISC IV		Raven	Math sub-tests average Z-scores and standard deviations (in brackets)										
		IQ	Z-score	Z-score	C	R	W	R	M	MC	SC	S	O	AP	TOT
Dyscalculics (N = 18)	12.1 (1.83) [8–15]	V: 107 (11) P: 100 (12) M: 74 (8) S: 89 (9) TOT: 94 (9)	V: −.3 (2) P: .1 (.8) M: −1.7 (1.6) S: −.7 (.6) TOT: −.4 (.6)	n/a	−1.7 (1.3)	−2.3 (1.5)	−1.9 (1.7)	−1.5 (1.3)	−3.6 (2)	−2 (1)	−2 (.6)	−1.6 (1.7)	−1.9 (1.7)	−2.3 (.8)	−2.1 (.7)
Controls (N = 32)	12.2 (.74) [10–13.5]	n/a	n/a	.6 (.8)	n/a	n/a	n/a	n/a	−.2 (.8)	.37 (.8)	−.3 (.9)	.36 (.3)	.1 (.6)	−.4 (.9)	0 (.5)

WISC IV= V: verbal reasoning, P: perceptual reasoning, W: working memory, S: speed processing.
Dyscalculia battery (BDE-2) = C: counting; R: repetition; W: writing; R:repetition; M: multiplication; MC: mental calculation; SC: speeded mental calculation; S: semantic; O: ordering; AP: approximation.
n/a: not available.

perceive and reproduce area and density independently we expect the two errors to be statistically independent and show little covariance. However, correlations are .49 on average and higher than .32 for all participants, all significant (see Table 2).

In our previous research with three-alternative forced choice methods, we found that at high densities the confusion ellipse became less elongated, and oriented more along the density than the numerosity axis. We therefore also measured reproduction with a denser comparison stimulus, with 128 dots within the same area (trials intermingled with the low density stimuli). The results, shown in Fig. 4, were quite different. Here the ovals are oriented closer to 0 than to 45° (mean = 9.6°, SD = 8°, see Table 2) and the correlations become weaker, .25 on average (see Table 2).

Fig. 5A, B plot the response of the "aggregate observer", the combination of all nine typically developing adults after removing individual average biases (centring each on the origin of the area/density plot) for low and high densities. For these distributions we calculated the error in the aggregate observer in all directions over a range of 360°. We plot these as overlay on error distributions as well as function of angle in Fig. 5C, D. This assumption free analysis confirms that for low density the direction with the lowest dispersion is close to the numerosity axis (40 and −148°) while for the dense patterns it is close to the area axis (3° and −175°).

One of the more influential recent hypotheses of numerosity perception is that numerosity is derived from independent estimates of patch area and density (Dakin et al., 2011; Durgin, 2008; Leibovich et al., 2016). If this were so, the error for numerosity estimate should be given by the sum of those for density and area (by propagation of errors): errors at other angles should be given by the weighted sum $\sigma_{AD} = \sqrt{\sigma_A^2 + \sigma_D^2}$. This model (thin pink lines) predicts a graded performance where the error of a combined measure is never less than the lowest of density and area, and is the same for the numerosity (45 and −135°) and the constant-numerosity axes (−45 and + 135°). This is clearly far from what was observed, with best performance along the numerosity axis, and worst when numerosity was constant. The shaded area between the two curves show the very large prediction errors, leading to a coefficient of determination $R^2 = .13$.

At high densities the results were quite different. The ellipse is oriented near 0°, with best sensitivity for area and worst for density. Here the prediction for all other thresholds from those for density and area is not too far from those obtained, and clearly follows the general pattern of the results: $R^2 = .97$.

3.2. *Typical and dyscalculic adolescents*

One open question is whether subjects diagnosed with math-learning disability (dyscalculia) have an impaired sense of number, as suggested by the fact that dyscalculics perform worse than age-matched controls in numerosity estimation and comparison tasks (Piazza, 2010; Piazza et al., 2010; Mazzocco et al., 2011). We recruited 18 adolescent dyscalculics and 35 age-matched typical subjects and assessed numerosity perception with both a classical explicit numerosity comparison task and with the new implicit reproduction task.

Table 2 – Summary statistics for psychophysical numerosity tasks.

Task and parameters	Pre-adolescent		Adults	
	Dyscalculics (N dots 12)	Typicals (N dots 12)	Typicals (N dots 12)	Typicals (N dots 128)
Reproduction, Angle	M = 40.807 STD = 8.893	M = 33.578 STD = 12.807	M = 36.04 STD = 10.793	M = 9.64 STD = 8
Reproduction, Short axis width (oct)	M = .240 STD = .119	M = .234 STD = .120	M = .13 STD = .03	M = .08 STD = .03
Reproduction, Response time (secs)	M = 10.13 STD = 5.32	M = 6.96 STD = 2.13	M = 5.61 STD = 1.19	M = 5.47 STD = 1.32
Reproduction, Area *vs* Density correlation	M = −.56 STD = .22	M = −.63 STD = .17	M = −.49 STD = .11	M = −.25 STD = .20
Comparison, Thresholds (oct)	M = .517 STD = .183	M = .846 STD = .42	na	na

Thresholds measured by explicit numerosity comparison task replicated previous studies showing that numerosity thresholds are higher for dyscalculics than for aged-matched controls (Fig. 8A). Mean weber fractions for the dyscalculic group were .74, compared with .49 for controls, significantly higher [one-tailed *t*-test, t(47) = 3.11, $p = .001$].

We then tested dyscalculic participants with our reproduction task. Fig. 6 displays error clouds for 18 dyscalculics tested with a low-density stimulus comprising 12 dots, and Fig. 7 shows the aggregate observer for the two subject groups. There was considerable variability in the orientation of the ellipses in the dyscalculic group, with several participants lying close to the numerosity axis but others closer to the area axis. The average ellipse orientation for the dyscalculic group was 33.5°, SD 13° (see also Fig. 8E and Table 2). This is closer to the area axis than for the typically developing participants (mean angle = 40.8°, SD = 9°, Table 2) (Fig. 8E). The difference between the two groups was significant [t(47) = 2.32, $p = .012$ one-tail].

We then looked at reproduction precision in three possible ways. The first is the width of the error clouds along the short axis (maximal sensitivity). This value gives an indication of the subject's precision along the axis that is most spontaneous for him or her. The two groups perform very similarly in this measure with average width of .24 octaves and .23 octaves for typicals and dyscalculics [t(47) = .28, $p = .6$ one-tail, Fig. 8B and Table 2]. The second error measure is the average Scatter defined as the scatter of responses around the average. This measure reflects both the scatter along the preferred direction as well as the scatter along all the other dimensions. Also this measure is very similar between groups [1.67 and 1.53 octaves for typicals and dyscalculics, t(47) = .85, $p = .8$ one-tailed *t*-test, Fig. 8C]. Lastly we considered total average error, which is the scatter of responses from the sample pattern, and therefore also includes biasing errors. Also this measure was very similar for the two groups [2.06 octaves for typicals and 2.08 for dyscalculics, t(47) = .14, $p = .44$ see Fig. 8D].

Although participants were not required to make a speeded response, we analysed average response time. Fig. 8F (see also Table 2 for summary statistics) shows that dyscalculics responded more slowly than typicals, with average reproduction times of 10.1 sec (SD = 5 sec), 40% higher than those of typicals (mean = 6.9 sec, SD = 2.1 sec). This difference is highly significant [t-test on log of RT: t(47) = 2.78, $p = .003$ one-tail], and remains significant even after removing outlier performance of one dyscalculic observer [t(46) = 2.37, $p = .01$].

For those perceptual parameters where dyscalculics were significantly different from typical participants, we looked for a correlation with math scores. Fig. 9 shows that both reproduction response times and thresholds in the 2AFC correlate significantly with math abilities (both $r > .3$ and $p < .05$). Interestingly, while the response time and discrimination thresholds correlate with math they did not correlate with each other ($r = .1$, $p = .48$), suggesting that they independently contribute in explaining math variance. We therefore ran two separate hierarchical regression models, with math ability as the dependent variable and one of the two numerosity performance measures as the independent variable with the remaining one as covariate. When controlling for reproduction response time, discrimination thresholds still explain a significant portion of math variance (R^2 change = 7.2%, $p = .048$). The same occurred for the response time effect after regressing out the influence of discrimination thresholds (R^2 change = 10.4%, $p = .019$).

We also looked for implicit learning effects during the reproduction experimental sessions (Fig. 10). For each subject, we divided the data into three blocks of 12 trials each and computed the average response time and short axis width for each time bin. Interestingly while Response Times of typicals decrease during the testing session (average slope = −.8 sec every block; Bootstrap $p = .018$), dyscalculics showed no significant change as a function of practise (slope −.3, $p = .76$). No significant effect on short axis width was found in either group.

One of the strengths of this technique is that it makes no explicit mention of numerosity, and can therefore also be employed in younger populations, pre-school and early primary school. To test feasibility in young children, we tested 3 pre-schoolers. All subjects found the task quite intuitive and enjoyable. Fig. 11 plots response distributions for sparse pattern (N12). While preliminary these data show that pre-schoolers show hallmarks of spontaneous numerosity processing, with the alignment of responses close to the diagonal, similarly to adults. Encouraged by these results we attempted to test a 3-year-old boy, but it was difficult to keep him focused on the task with this version. However, we are confident that introducing a more child-friendly version, this technique could be extended to younger participants.

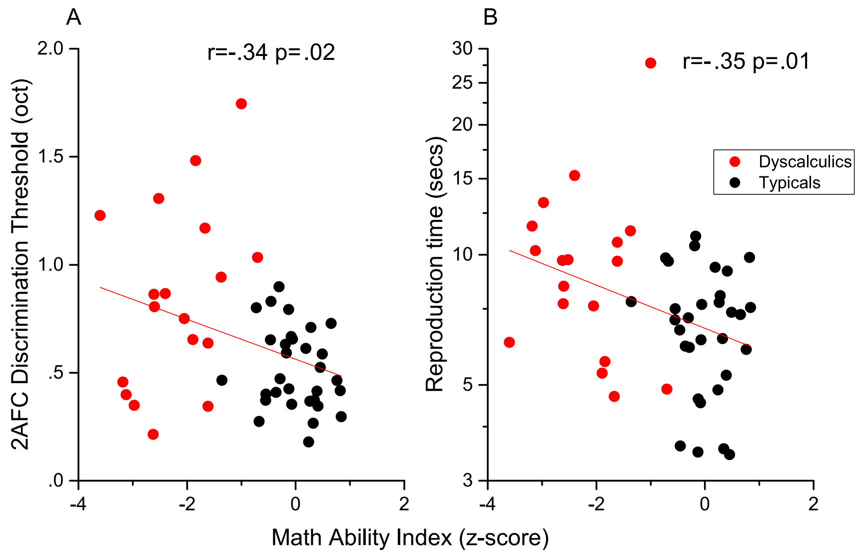

Fig. 9 – Correlation between perceptual scores and math ability. For each individual we plot performance in the perceptual tasks (2AFC Weber fraction or Reproduction time) against average scores in a battery of mathematical tests. Black indicates typical adolescents, red, dyscalculic adolescents.

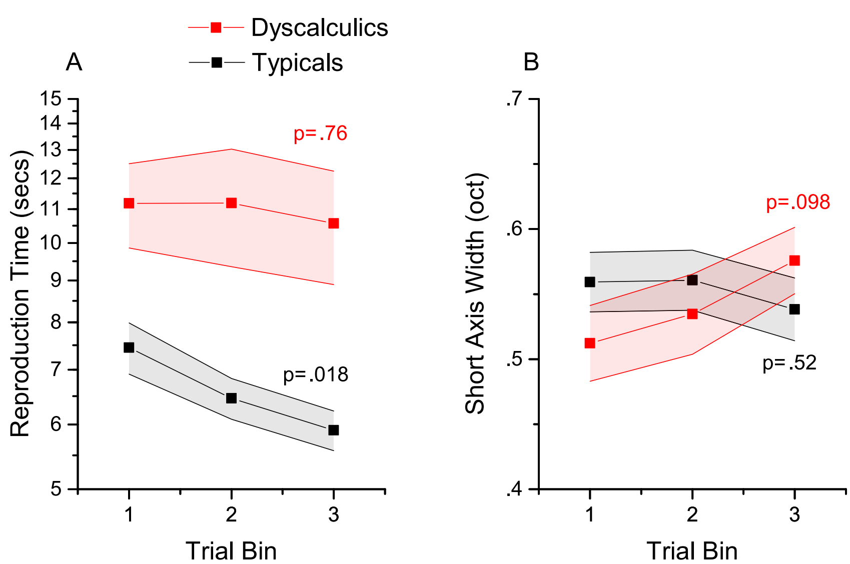

Fig. 10 – Learning effects in the reproduction task. A) Average response time during the experiment considering three bins of 12 trials each. B) Average short axis width as function of time. Black typical adolescents, Red dyscalculics adolescents. Numbers in the inset indicate two-tail bootstrap *t*-test of linear fit slope against zero (10,000 iterations).

3.3. *Task reliability*

Reliability measures are usually obtained by test retest techniques (Portney & Watkins, 1993). However, this reduces the data by half (particularly problematic for psychometric functions), as well as introducing sequence effects. We therefore measured reliability with a bootstrap technique that circumvents both these problems, and has been shown to be useful in development studies (Anobile, Castaldi, et al., 2016; Anobile, Arrighi, Castaldi, Grassi, Pedonese, & PA, 2017). We ran a Montecarlo simulation by correlating across subject responses the estimates of width, reaction time and response angle taken from two random samples (with as many trials as original data, sampled with replacement). We run the simulation 10,000 times, to yield 10,000 separate estimates of correlation and then computed a grand average (reliability level). For the reproduction task, mean correlations were: .68 for the short axis parameter; .91 for reaction times; and .53 for the reproduction angle. We used the same technique to measure reliability of the Weber fractions obtained by 2 AFC discrimination, and found an average correlation of .75, suggesting that the reliability levels of the two technique are comparable.

4. Discussion

In this study we present a novel technique to measure numerosity, area and density encoding, by asking subjects to

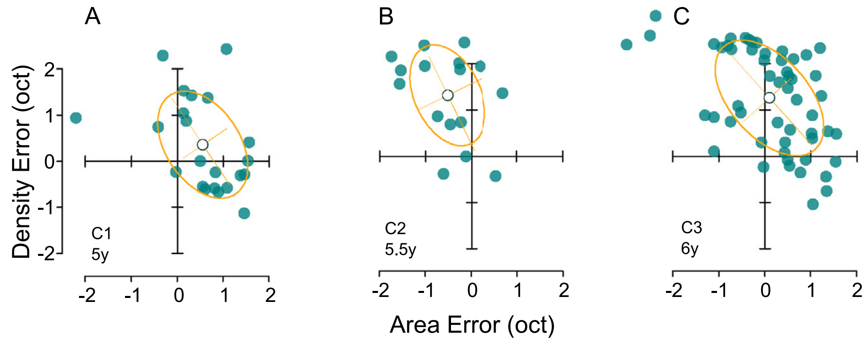

Fig. 11 – Performance of 5 yr old children for low density stimuli A–C) Errors in the reproduction task of three typical developing pre-schooler children aged between 5 and 6 years at numerosity 12. Conventions as before.

reproduce with a two-dimensional trackpad or mouse a briefly displayed cloud of dots. The task is easy, intuitive and quite enjoyable. Most importantly, as subjects were not given explicit instructions of what to match, it captures which feature is spontaneously perceived. We found the task useful not only with typically developing adults, but also with adolescents and subjects with math-learning disorder (dyscalculia).

The results of this study support and extend those obtained with rigorous forced-choice techniques, shown in Fig. 1 (Cicchini et al., 2016). For most typical adult observers, for the low pattern density, errors fall along an ellipse oriented near the constant numerosity axis. The average orientation of the ellipse was 36° (whereas the numerosity axis is 45°), suggesting that the confusion ellipse is dominated 2/3 by numerosity and 1/3 by area. We also replicate that the confusion ellipses change orientation for dense patterns in typical adult participants, aligning more closely to the area than the numerosity axis.

It has been proposed that numerosity may not be sensed directly, but derived indirectly from density mechanisms (Dakin et al., 2011; Durgin, 2008) or integrating multiple visual cues (Gebuis et al., 2016). The results with sparse displays clearly run contrary to this idea, as mechanisms that solve the matching task considering solely area and density would have error clouds aligned to the major axes. Fig. 5C, D plot the width of the error function for the aggregate subject (after alignment), as a function of angle. For sparse patterns, the width is least (implying maximum sensitivity) near the numerosity axes at (45 and −135°), and greatest near the axis of constant numerosity (−45 and + 135°). Simulation showed that the data could not be explained by the independent assessment of density and area. On the other hand, at the higher density of 128 dots, where we suggest that texture-density mechanisms operate (Anobile et al., 2014; Cicchini et al., 2016; Anobile, Cicchini, et al., 2017), the data were well explained by a simple model with area and density as cardinal axes, with errors at other orientations resulting from a weighted sum of the errors at these orientations. Our observations are further backed up by the recent findings that numerosity and size estimation have different developmental trajectories, and that precision and adaptation strength of the two systems do not correlate (Anobile, Burr, Iaia, Marinelli, Angelelli, & Turi, 2018; Anobile, Cicchini, Gasperini, & Burr, 2018).

The technique provides a promising potential tool to study development of numerosity, and also disorders. A typical test lasts only 4.5 min, while the three-alternative forced-choice of Cicchini et al. (2016) takes 130 min for a two-dimensional psychometric function. Even a simple Weber fraction, calculated by two-alternative forced-choice takes up to than 7–8 min. Despite the ease of use and brief testing time, the reaction times split-half reliability level were similar to those of traditional measures, and comparable to that for measuring Weber fractions by standard two-alternative forced-choice techniques. Note that this technique could be simplified further to vary only numerosity (keeping area fixed), if that were the only variable of interest.

Given the potential of the technique in the clinic and the growing interest in number sense development in dyscalculia, we used it to measure emergence of numerosity sensitivity in typical and dyscalculic adolescents. We chose to study this age group, as diagnosis at this developmental stage is likely to reflect a genuine neurodevelopmental disorder in number processing, rather than a temporary difficulty during early math acquisition, where the intersubjective variability is very high even in the typical development. Surprisingly, dyscalculics did not show higher response scatter in the array reproduction task, something that would be quite expected given their lower precision in explicit numerosity comparison (replicated in this study). On a face value this may suggest that the precision in the numerosity representation in dyscalculics is similar to controls. However, they also exhibited changes in two other related measures, which may themselves indicate a difficulty in processing number. One was a small tilt of average responses towards the area axis, which suggests that dyscalculics do not employ only numerosity but also use area information. This may reflect increased difficulty in handling numerosity information, so they supplement numerosity with other information, keeping overall error relatively low.

Secondly, reaction times exceeded those of controls by about 40%. Slower reproduction times could reflect a genuine deficit in the processing speed, but could also indicate that compensatory mechanisms occur, and these add to the response times. Consistently, in the low-math ability group we found that higher response times were associated with more precise reproductions ($r = -.42$ $p = .07$) and with ellipses leaning towards the area axis ($r = -.2$, $p = .4$). Overall this suggests that those who cannot manipulate correctly number information, and thus take more time, also resort to stimulus dimensions other than numerosity, such as area and obtain a scatter of responses similar to that of controls. Interestingly, Piazza et al. (2010) have shown that in an explicit numerosity discrimination task, where subjects are allowed to inspect the dot-clouds as much as needed, dyscalculics perform as rapidly as age matched typicals but have a strong impairment in the precision level (thresholds). All these results suggest that different numerosity tasks requirements may impinge on different impairments of number processing in dyscalculia.

Finally, we also found that in typical subjects, response times decreased throughout the experimental session, whereas for dyscalculics they were constant throughout data collection. This is reminiscent of the learning disabilities that they display with symbolic math: but would requires further studies to confirm.

Overall our study has demonstrated that signatures of spontaneous numerosity processing can be studied with a very fast array reproduction task, without need for explicit mention of numerosity. This new technique has therefore the important advantage of not requiring any verbal instruction and it thus reflect a tool well suitable for investigation in critical populations, such as pre-school children and clinical populations.

Acknowledgments

This research was funded by the Italian Ministry of University and Research under the project "Futuro in Ricerca", Grant number RBFR1332DJ, and by the European Research Council FP7-IDEAS-ERC- under the project "Early Sensory Cortex Plasticity and Adaptability in Human Adults – ECSPLAIN-" Grant number 338866; and from Italian Ministry of Health and by Tuscany Region under the project "Ricerca Finalizzata", Grant n. GR-2013-02358262 to GA. We thank all the families of the children involved in the study and the Neuropsychology Unit of IRCCSS Stella Maris (Calambrone) in particular Dr Anna Maria Chilosi and Dr Filippo Gasperini for recruitment of dyscalculic participants.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2018.11.019.

REFERENCES

Aagten-Murphy, D., & Burr, D. (2016). Adaptation to numerosity requires only brief exposures, and is determined by number of events, not exposure duration. *Journal of Vision, 16*(10), 22. https://doi.org/10.1167/16.10.22.

Anobile, G., Arrighi, R., Castaldi, E., Grassi, E., Pedonese, L., PA, M. M., et al. (2017). Spatial but not temporal numerosity thresholds correlate with formal math skills in children. *Developmental Psychology*. https://doi.org/10.1037/dev0000448.

Anobile, G., Arrighi, R., Togoli, I., & Burr, D. C. (2016). A shared numerical representation for action and perception. *Elife, 5*. https://doi.org/10.7554/eLife.16161.

Anobile, G., Burr, D. C., Iaia, M., Marinelli, C. V., Angelelli, P., & Turi, M. (2018). Independent adaptation mechanisms for numerosity and size perception provide evidence against a common sense of magnitude. *Scientific Reports, 8*(1), 13571. https://doi.org/10.1038/s41598-018-31893-6.

Anobile, G., Castaldi, E., Turi, M., Tinelli, F., & Burr, D. C. (2016b). Numerosity but not texture-density discrimination correlates with math ability in children [Article] *Developmental Psychology, 52*(8), 1206–1216. https://doi.org/10.1037/dev0000155.

Anobile, G., Cicchini, G. M., & Burr, D. C. (2014). Separate mechanisms for perception of numerosity and density. *Psychological Science, 25*(1), 265–270. https://doi.org/10.1177/0956797613501520.

Anobile, G., Cicchini, G. M., & Burr, D. C. (2016). Number as a primary perceptual attribute: A review [Review] *Perception, 45*(1–2), 5–31. https://doi.org/10.1177/0301006615602599.

Anobile, G., Cicchini, G. M., Gasperini, F., & Burr, D. C. (2018). Typical numerosity adaptation despite selectively impaired number acuity in dyscalculia. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2018.10.006.

Anobile, G., Cicchini, G. M., Pomè, A., & Burr, D. C. (2017). Connecting visual objects reduces perceived numerosity and density for sparse but not dense patterns. *Journal of Numerical Cognition*. https://doi.org/10.5964/jnc.v3i2.38.

Anobile, G., Turi, M., Cicchini, G. M., & Burr, D. C. (2012). The effects of cross-sensory attentional demand on subitizing and on mapping number onto space. *Vision Research, 74*, 102–109. https://doi.org/10.1016/j.visres.2012.06.005.

Anobile, G., Turi, M., Cicchini, G. M., & Burr, D. C. (2015). Mechanisms for perception of numerosity or texture-density are governed by crowding-like effects. *Journal of Vision, 15*(5), 4. https://doi.org/10.1167/15.5.4.

Arrighi, R., Togoli, I., & Burr, D. C. (2014). A generalized sense of number. *Proceedings of the Royal Society of London B Biological Sciences, 281*(1797). https://doi.org/10.1098/rspb.2014.1791.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436. https://doi.org/10.1163/156856897x00357.

Burr, D. C., Anobile, G., & Arrighi, R. (2017). Psychophysical evidence for the number sense. *Philosophical Transactions of the Royal Society B Biological Sciences, 373*(1740). https://doi.org/10.1098/rstb.2017.0045.

Burr, D. C., Anobile, G., & Turi, M. (2011). Adaptation affects both high and low (subitized) numbers under conditions of high attentional load. *Seeing and Perceiving, 24*(2), 141–150. https://doi.org/10.1163/187847511X570097.

Burr, D. C., & Ross, J. (2008). A visual sense of number. *Current Biology, 18*(6), 425–428. https://doi.org/10.1016/J.Cub.2008.02.052.

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision, 10*(6). https://doi.org/10.1167/10.6.20.

Butterworth, B. (1999). *The mathematical brain*. London: Macmillan.

Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: From brain to education. *Science, 332*(6033), 1049–1053. https://doi.org/10.1126/science.1201536.

Cicchini, G. M., Anobile, G., & Burr, D. C. (2016). Spontaneous perception of numerosity in humans. *Nature Communications, 7*, 12536. https://doi.org/10.1038/ncomms12536.

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 of the United States of America, 108*(49), 19552–19557. https://doi.org/10.1073/Pnas.1113195108.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*(1–2), 1–42. https://doi.org/10.1016/0010-0277(92)90049-N.

Dehaene, S. (2011). *The number sense : How the mind creates mathematics (Rev. and updated ed.)*. New York: Oxford University Press.

Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Current Biology, 18*(18), R855–R856. https://doi.org/10.1016/J.Cub.2008.07.053.

Egeth, H. E., Leonard, C. J., & Palomares, M. (2008). The role of attention in subitizing: Is the magical number 1? *Visual Cognition, 16*(4), 463–473. https://doi.org/10.1080/13506280801937939.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002.

Ferrigno, S., Jara-Ettinger, J., Piantadosi, S. T., & Cantlon, J. F. (2017). Universal and uniquely human factors in spontaneous number perception. *Nature Communications, 8*, 13968. https://doi.org/10.1038/ncomms13968.

Gebuis, T., Cohen Kadosh, R., & Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: A critical review. *Acta Psychologica, 171*, 17–35. https://doi.org/10.1016/j.actpsy.2016.09.003.

Gebuis, T., & Reynvoet, B. (2012). The role of visual information in numerosity estimation. *Plos One, 7*(5), e37426. https://doi.org/10.1371/journal.pone.0037426.

Harvey, B. M., Klein, B. P., Petridou, N., & Dumoulin, S. O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science, 341*(6150), 1123–1126. https://doi.org/10.1126/science.1239052.

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences, 106*(25), 10382–10385. https://doi.org/10.1073/pnas.0812142106.

Jevons, W. S. (1871). The power of numerical discrim- ination. *Nature, 3*, 363–372.

Kaufman, E. L., & Lord, M. W. (1949). The discrimination of visual number. *The American Journal of Psychology, 62*(4), 498–525.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception, 36*, 14–14.

Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2016). From 'sense of number' to 'sense of magnitude' - the role of continuous magnitudes in numerical cognition. *The Behavioral and Brain Sciences*, 1–62. https://doi.org/10.1017/S0140525X16000960.

Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development, 82*(4), 1224–1237. https://doi.org/10.1111/j.1467-8624.2011.01608.x.

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 of London B Biological Sciences, 281*(1790). https://doi.org/10.1098/rspb.2014.1137.

Odic, D. (2017). Children's intuitive sense of number develops independently of their perception of area, density, length, and time. *Developmental Science*. https://doi.org/10.1111/desc.12533.

Olivers, C. N. L., & Watson, D. G. (2008). Subitizing requires attention. *Visual Cognition, 16*(4), 439–462. https://doi.org/10.1080/13506280701825861.

Park, J., DeWind, N. K., Woldorff, M. G., & Brannon, E. M. (2016). Rapid and direct encoding of numerosity in the visual stream. *Cerebral Cortex, 26*(2), 748–763. https://doi.org/10.1093/cercor/bhv017.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442. https://doi.org/10.1163/156856897x00366.

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences, 14*(12), 542–551. https://doi.org/10.1016/j.tics.2010.09.008.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., et al. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition, 116*(1), 33–41. https://doi.org/10.1016/j.cognition.2010.03.012.

Pinheiro-Chagas, P., Wood, G., Knops, A., Krinzinger, H., Lonnemann, J., Starling-Alves, I., et al. (2014). In how many ways is the approximate number system associated with exact calculation? *Plos One, 9*(11), e111155. https://doi.org/10.1371/journal.pone.0111155.

Portney, L. G., & Watkins, M. P. (1993). *Foundations of clinical research: Applications to practice*. Norwalk, Conn.: Appleton & Lange.

Railo, H., Koivisto, M., Revonsuo, A., & Hannula, M. M. (2008). The role of attention in subitizing. *Cognition, 107*(1), 82–104. https://doi.org/10.1016/j.cognition.2007.08.004.

Ross, J. (2003). Visual discrimination of number without counting. *Perception, 32*(7), 867–870. https://doi.org/10.1068/P5029.

Ross, J., & Burr, D. C. (2010). Vision senses number directly. *Journal of Vision, 10*(2). https://doi.org/10.1167/10.2.10, 10 11-18.

Stoianov, I., & Zorzi, M. (2012). Emergence of a 'visual number sense' in hierarchical generative models. *Nature Neuroscience, 15*(2), 194–196. https://doi.org/10.1038/nn.2996.

Tibber, M. S., Greenwood, J. A., & Dakin, S. C. (2012). Number and density discrimination rely on a common metric: Similar psychophysical effects of size, contrast, and divided attention. *Journal of Vision, 12*(6). https://doi.org/10.1167/12.6.8.

Tibber, M. S., Manasseh, G. S. L., Clarke, R. C., Gagin, G., Swanbeck, S. N., Butterworth, B., et al. (2013). Sensitivity to numerosity is not a unique visuospatial psychophysical predictor of mathematical ability. *Vision Research, 89*, 1–9. https://doi.org/10.1016/J.Visres.2013.06.006.

Vetter, P., Butterworth, B., & Bahrami, B. (2008). Modulating attentional load affects numerosity estimation: Evidence against a pre-attentive subitizing mechanism. *Plos One, 3*(9). https://doi.org/10.1371/journal.pone.0003269.

Watson, A. B., & Pelli, D. (1983). Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113–120.