# Simultaneous and sequential subitizing are separate systems, and neither predicts math abilities

# Simultaneous and sequential subitizing are separate systems, and neither predicts math abilities.

Giovanni Anobile[1], Roberto Arrighi[2], David C. Burr[2,3,4]

1. Department of Developmental Neuroscience, Stella Maris Scientific Institute, Calambrone Pisa, Italy
2. Department of Neuroscience, Psychology, Pharmacology and Child Health, University of Florence, Florence, Italy
3. School of Psychology, University of Sydney, Sydney, Australia
4. Department of Translational Research on New Technologies in Medicines and Surgery, University of Pisa, 56123 Pisa, Italy

Keywords: Numerosity, Subitizing, Numerical cognition, Approximate Number System, Dyscalculia

Corresponding author:
Dott. Giovanni Anobile (Phd)
e-mail: giovannianobile@hotmail.it
Department of Developmental Neuroscience, Stella Maris Scientific Institute, Viale del Tirreno,
331, 56018 Calambrone (Pisa), Italy

**Acknowledgments.** This research was funded by the Italian Ministry of Health and by Region of Tuscany under the project "*Ricerca Finalizzata*", Grant n. GR-2013-02358262 to GA; 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. GA was partially founded by FP7-IDEAS-ERC- grant number 338866.

# Simultaneous and sequential subitizing are separate systems, and neither predicts math abilities

**Keywords:** Numerosity, Subitizing, Numerical cognition, Approximate Number System, Dyscalculia

**ABSTRACT**

Small quantities of visual objects can be rapidly estimated without error, a phenomenon known as subitizing. Larger quantities can also be rapidly estimated, but with error: and the error rate predicts math abilities. This study addresses two issues: does subitizing generalize over modalities and stimulus formats? – and does subitizing correlate with math abilities? We measured subitizing limits in primary school children and adults, for visual and auditory stimuli presented either sequentially (sequences of flashes or sounds), and for simultaneously visual presentations (dot arrays). The results show: a) Subitizing limits for adults were one item larger than for primary school children across all conditions; b) Subitizing for simultaneous visual stimuli (dots) was better than that for sequential stimuli; c) Subitizing limits for dots do not correlate with subitizing limits for either flashes or sounds; d) Subitizing of sequences of flashes and sounds are strongly correlated with each other in children; e) Regardless of stimuli sensory modality and format, subitizing limits do not correlate with mental calculation or digit magnitude knowledge proficiency. These results suggest that although children can subitize sequential numerosity,

simultaneous and temporal subitizing may be sub-served by separate systems. Furthermore, subitizing does not seem to be related to numerical abilities.

## INTRODUCTION

Although humans are the only species to have evolved a symbolic language-based code for mathematical concepts, we share with many animals the ability to make rough but rapid estimates of object numerosity (Agrillo, Miletto Petrazzini, & Bisazza, 2017; Dacke & Srinivasan, 2008; Dehaene, 2011; Ditz & Nieder, 2015; Nieder, 2016; Petrazzini, Agrillo, Izard, & Bisazza, 2016; Rugani, Vallortigara, Priftis, & Regolin, 2015). In general, numerosity estimates are fast and errorless up to 4-6 items; after this range, performance decreases monotonically, both reaction times (RTs) and accuracy. Kaufman (1949) was first to coin the term "subitizing", derived from the Latin *subitus* meaning sudden. Subitizing can be defined in several ways. Kaufman (1949) measured the subitizing limit as the point of discontinuity in the distribution of RTs or accuracy. This typically resulted in subitizing being defined as occurring for stimulus numerosities below 6. Subsequent studies employed slightly different definitions, all based on performance discontinuities and resulting in slightly different estimates of the subitizing limit (Arp & Fagard, 2005; Arp, Taranne, & Fagard, 2006; Ashkenazi, Mark-Zigdon, & Henik, 2013; Burr, Turi, & Anobile, 2010; Camos & Tillmann, 2008; Green & Bavelier, 2003; Olivers & Watson, 2008; Piazza, Fumarola, Chinello, & Melcher, 2011; Revkin, Piazza, Izard, Cohen, & Dehaene, 2008; Schleifer & Landerl, 2011). One way is to take the inflexion point of a sigmoid function fitting estimation's error or reaction times (Piazza et al., 2011; Revkin et al., 2008). This method tends to overestimate the limit compared with other techniques, but has been proven to be robust, and well suited to detect inter-individual differences.

Subitizing phenomena are thought to be linked to many non-numerical capacities like attention, working memory and object tracking. All these mechanisms are capacity-limited and interact with each other, making it difficult to assess their individual contribution. For example, simultaneous subitizing is heavily affected by the deployment of simultaneous and temporal attentional resources (Burr et al., 2010; Olivers & Watson, 2008; Vetter, Butterworth, & Bahrami, 2008) as well as visual working memory (Melcher & Piazza, 2011; Piazza et al., 2011). All these results clearly suggest the existence of partially shared mechanisms. The so-called "Object Tracking System" (OTS), the process involved in identifying, representing and tracking objects through time and space is, like subitizing, strongly dependent on attentional resources (Arrighi, Lunardi, & Burr, 2011; Pylyshyn & Storm, 1988). However, there are clear differences between these processes. For example, OTS capacity has been found to be adult-like at one year of age (Piazza, 2010), while spatial working memory continues to develop until 6/7 years (Cowan, Morey, AuBuchon, Zwilling, & Gilchrist, 2010). Also OTS capacity measured by a visual multiple object tracking task is affected by visual but not auditory attentional deprivation (Arrighi et al., 2011), while visual subitizing strongly suffers from cross-modal (visual, auditory and haptic) dual tasks (Anobile, Turi, Cicchini, & Burr, 2012). Some evidence also suggests a link between individual differences in working memory (Bull & Scerif, 2001; De Smedt et al., 2009; Passolunghi & Siegel, 2004; Raghubara, Barnesb, & Hechtb, 2010; Toll, Kroesbergen, & Van Luit, 2016), attention (Ashkenazi & Henik, 2012; Askenazi & Henik, 2010; Steele, Karmiloff-Smith, Cornish, & Scerif, 2012) as well as OTS (Anobile, Stievano, & Burr, 2013) and math performance. In brief, while subitizing has been extensively studied, the underlying mechanisms, and how it relates to other cognitive capacities are still unclear.

Moreover, while subitizing has been extensively studied, most studies concentrate on spatial arrays of simultaneous visual stimuli; and it is still not clear whether subitizing generalizes to other sensory modalities. Very few studies have investigated subitizing for temporal sequences, or in sensory modalities other than vision. The available data suggest that adults may subitize auditory sequences (Camos & Tillmann, 2008; Repp, 2007), but not simultaneously played sounds (McLachlan, Marco, & Wilson, 2012). In vision, only one study has described subitizing for temporal sequences (Camos & Tillmann, 2008), while more evidence has been provided for the existence of subitizing sequences of haptic stimuli (Ferrand, Riggs, & Castronovo, 2010; Gallace, Tan, Haggard, & Spence, 2008; Plaisier, Bergmann Tiest, & Kappers, 2009; Plaisier & Smeets, 2011; Plaisier, Tiest, & Kappers, 2010; Riggs et al., 2006). All these studies involved adult participants, and none took into account between-task correlations.

It has been suggested that subitizing may also be fundamental for learning more complex numerical processes. For example, Carey (2002) proposed that subitizing and OTS are ideal to represent natural numbers, and thus provide the first meaning of numerals to children. In line with this theory, a study from our group showed that in primary school children, OTS capacity measured by multiple object tracking, positively correlates with math abilities (Anobile et al., 2013). However, the tracking task requires tracking objects in space and time, and might not tap on the same perceptual process involved in enumeration of simultaneous and rapidly presented arrays. Moreover, dyscalculic subjects do not show clear peculiarities for either OTS or subitizing (Piazza, 2010). One study involving dyscalculic adolescents reported almost typical subitizing capacity (Ceulemans et al., 2014), while another reported that from 43% and 79% dyscalculic subjects in the age range of 7–17 years had impaired subitizing, more evident in older subjects (Fischer, Gebhardt, & Hartnegg, 2008). In line with this, only the 30% of dyscalculic children around

8.5 years old showed subitizing difficulties (Desoete & Grégoire, 2006). In summary, the evidence for a link between subitizing and math is somewhat variable.

The numerical range above subitizing is termed the “estimation” range, thought to reflect the action of the “Approximate Number System (ANS)” (Feigenson, Dehaene, & Spelke, 2004). The ANS is a generalized system, encoding and integrating information across sensorimotor domains, including vision, audition and action, and stimulus formats, both simultaneous spatial ensembles and temporal sequences (Anobile, Arrighi, Togoli, & Burr, 2016; Arrighi, Togoli, & Burr, 2014; Izard, Sann, Spelke, & Streri, 2009). ANS precision correlates with current and future children math abilities along the entire spectrum of mathematical abilities, from low-math to math-gifted children (Piazza et al., 2010; Wang, Halberda, & Feigenson, 2017), including ‘average’ math skilled children.

The goal of this study is to investigate subitizing, testing for generalization across format and modality. We will study its developmental trend, and look for correlations with formal mathematical skills. The results show a clear developmental trend in subitizing capacity, but no relationship between simultaneous and temporal subitizing, or between subitizing and math.

## METHODS

**Participants.** 98 children (7.1 – 11.0 years old, mean 9.2), and 38 adults (19 – 30 years, mean 25.6) were included in this study. Children were recruited from local schools, and only those who returned a signed consent from parents were included. Experimental procedures were approved by the local ethics committee (*Comitato Etico Pediatrico Regionale—Azienda Ospedaliero-Universitaria Meyer*—Florence, Italy) and are in line with the declaration of Helsinki.

**General procedures.** Stimuli were generated and presented with Matlab 8.1 using PsychToolbox routines (Brainard, 1997) on a 17-inch LG touch screen monitor with 1280 X 1024 resolution at refresh rate of 60 Hz. Each participant was tested in two separate sessions (usually within the same week), lasting around 1 hour each. Math abilities were measured by a paper and pencil test (only children), and by a computerized digit summation task (children and adults). All participants also performed a non-verbal reasoning task (Raven matrices). Math skills and nonverbal reasoning were measured at the end of the first session, perceptual tasks were administered in a pseudorandom order between participants. This study is based on a new analysis of a set of data collected for other purposes (Anobile et al., 2017). The experimental methods used here were the same as in the previous study but here we focused on the subitizing range by not excluding numerosities ≤4 from the analyses.

**Numerosity estimation.** Visual stimuli were either ensembles of 0.5° diameter dots half-white, half-black in order to balance luminance across numerosities (in case of odd numbers the one excess dot was randomly assigned to white or black), presented simultaneously for 250 ms within a virtual 16° diameter region, or sequences of flashes (sharp-edged white discs of 90 cd m−2 and 5° diameter) presented in a pseudo-random order within a 2 second interval (Figure 1 A and B). In the sequential presentation, each flash lasted 40 ms with the constraint that two pulses could not fall within 40 ms of each other. All visual stimuli were presented centrally, with subject viewing distance set at 57 cm, on a grey background of 40 cd/$m^2$. Precision for estimates of sequential numerosity was also investigated in audition, with 500 Hz pure tones ramped on and off with 5 ms raised-cosine ramps, presented with an intensity of 80 dB (at the sound source) and digitized at a rate of 65 kHz. Sounds were presented through high-quality headphones Microsoft lifechat LX-3000, and perceptually localized in the middle of the head. In all conditions, the numerosity range was 2-18 and subjects were asked to verbally report the number of perceived stimuli, which was recorded by the experimenter via a computer keyboard. The testing phase was preceded by a training session of 17 trials (not included in the main analyses). During training, all numerosities were randomly presented, and feedback was provided displaying the actual numerosity on the monitor screen. The aim of feedback was to calibrate participants' judgements (mainly young children) to have all estimates within the numerical range without aberrant responses (for a similar procedure see Revkin et al., 2008). After training had been completed, the testing phase started with a block of 51 trials (3 repetitions for each numerosity), with no feedback. In total each participant performed 204 trials. Test numerosity ranged from 2 to 18, but we computed error rates and fitted (see later) only the range 2-16 to avoid edge effects. Average temporal rates for both flashes and sounds stimuli were 640 ms (min 110, max 1180), 500 ms (min 130, max 900) and 400 ms (min 140, max 650) respectively for numerosity two, three and four. As the counting speed for numbers in Italian primary school children is around 600/800 milliseconds per number, and the stimulus sequences were not regular but jittered in time, it is unlikely that children were able to serially count the stimuli.

**Semantic skills.** Two paper and pencil tasks were administered (Figure 1 C): 1) mark the largest numbers in a set of three (one to five digits, 36 trials); and 2) mark where a number should be placed (four possible positions among three other numbers, one to six digits, 18 trials). These tasks were extracted from an Italian standardized battery suitable for children from 8 to 13 years old, not suitable for adults (Biancardi, Bachmann, & Nicoletti, 2016). They are thought to tap the semantic component of numeracy (Dehaene, Piazza, Pinel, & Cohen, 2003), and have been demonstrated to be good predictor of children numerosity discrimination thresholds (Anobile et al., 2013; Cicchini, Anobile, & Burr, 2016; Piazza, 2010). Again, accuracy and speed were measured (as the sum of errors and time in minutes required to complete the three tasks). Similarly to the mental calculation task (see below) we measured two separate z-scores for speed and accuracy and computed a performance-combined index averaging the two z-scores (same technique exploited by Anobile et al., 2017; Cicchini et al., 2016).

**Computerized mental addition task.** On each trial three digits (3°×3°, Geneva font) were displayed, two (vertically aligned at a distance of 1.5°) to the left and one to the right of a central dot reference point (horizontal eccentricity 2°). We asked participants to mentally sum quickly but accurately the two digits on the left and compare the result with the single digit on the right (Figure 1 D). Responses were self-provided indicating (by appropriate key press) which side contained the higher magnitude. Both the addenda ranged from 1 to 9 and were randomly chosen, on each trial, with the sum of the two numbers constrained between 5 and 10 (grain of 1). The single digit (comparison sum) was determined by adding to the real sum a delta value chosen from a flat distribution ranging from ±60% for children, and ±40% for adults, rounding to the closest integer. Participants performed a total of 70 trials divided in two separate blocks of 35 each. We applied a time threshold (2 and 5 secs for adults and children respectively), with thresholds derived from pilot data. In trials where RTs exceeded the threshold 5.6% and 1.8% for children and adults respectively, we gave an auditory feedback. The feedback did not provide any information about the accuracy, only of the need to perform the operations more quickly.

Not every trial where RTs exceeded the threshold were eliminated from the analysis, as we applied a within subject cut-off: for each participant we measured the average reaction times (across trials) and eliminated those higher or lower than 3 standard deviations. The total number of eliminated trials was 38 (1.1%) for adults and 80 (1.4%) for children. The proportion of 'sum higher' was plotted against the percentage difference between the sum and the single digit. We fitted the data with a cumulative Gaussian error functions. The percentage difference needed to move from 50% to 75% correct responses provided an “mental additional discrimination threshold”. This is logically equivalent to the Weber fraction usually measured for numerosity discrimination tasks, and could be interpreted as the amount of noise present in the mental addition process (see Figure 1 D). Similarly to previous studies (Anobile et al., 2017; Cicchini et al., 2016), we computed for each participant two separate z-scores: one for precision (Weber fraction) and the other for response speed (RT). Z-scores were measured using the mean and standard deviation of the participant grade class (from second to fifth grade). For adults we used the mean and standard deviation of the entire group. Finally, for each participant we computed a performance-combined index averaging the two z-scores. A previous study demonstrated that children’s performance on this task is a good predictor of their numerosity estimation precision of simultaneous dot arrays (Weber Fraction), for numerosity above the subitizing range (Anobile et al., 2017).

**Data analysis.** Following previous works (Piazza et al., 2011; Revkin et al., 2008), we fitted error rates with sigmoid functions and defined the subitizing range as the inflection point of the function. We performed this procedure separately for each participant as well as for average data (lines on Figure 2). As noted by others, this procedure may overestimate the subitizing limit, but this should bias all conditions equally. On the other hand, the fitting procedure has proven to be very robust, particularly in capturing individual variability necessary for correlational studies (Piazza et al., 2011; Revkin et al., 2008).

Correlation analyses were performed by both zero-order and partial Pearson correlations procedures. Statistical significance was indexed by p-values and Bayes factor (Wetzels & Wagenmakers, 2012). Bayes factor is the ratio of the likelihood probabilities of the two models, that a correlation s quantifies the ratio of the likelihood probabilities $H_1/H_0$, where $H_1$ is the likelihood of a correlation between the two variables, and $H_0$ the likelihood that the correlation does not exist. By convention, a Log Bayes Factor (LBF) greater than 0.5 it is considered substantial evidence in favour of the existence of the correlation, and LBF < −0.5 substantial evidence in favour of it not existing. Absolute values of LBF greater than 1 are considered strong evidence, and values greater than 2 are considered decisive. Missing values were left empty and data excluded with pairwise deletion method.

**Task reliability.** We measured reliability using split-half bootstrap techniques. 1) Mental addition task. For each participant we calculated two separate thresholds (or RTs) from a random sample of the data (70 trials, as large as the data set taken, sampled with replacement from the data set), and then computed the correlation between those two measures (Pearson-r). We reiterated the process 1000 times for all participants, to yield mean and standard error estimate of reliability. This method is validated and described in Cicchini et al. (2016). 2) Subitizing limits. The R-squares of the fits were reasonably high, suggesting it was an appropriate measurement procedure (see results). However, we also measured two other indexes of reliability. The first analysis mirrors that described above, except that on each iteration for each participant we calculated two separate subitizing limits. As for the main analysis, we eliminated values with $R^2$ lower then 0.25 (10 % overall). In the second analysis we looked at pooled data: 1) for each numerosity we pooled together all the trials, 2) separately for each numerosity, we divided the trials into two equal-n samples by randomly sampling the data (half size than the data set taken, sampled with replacement from the full data set) 3) we fitted these two separate data sample with the procedure described above, producing two measurements of subitizing limits. On each reiteration (1000 times) we calculated the different in the limits of the two conditions, and counted the proportion of times one was higher than another (sign test). Table 1 reports average subitizing capacity measurements for the two data-halves with associated difference and p-values.

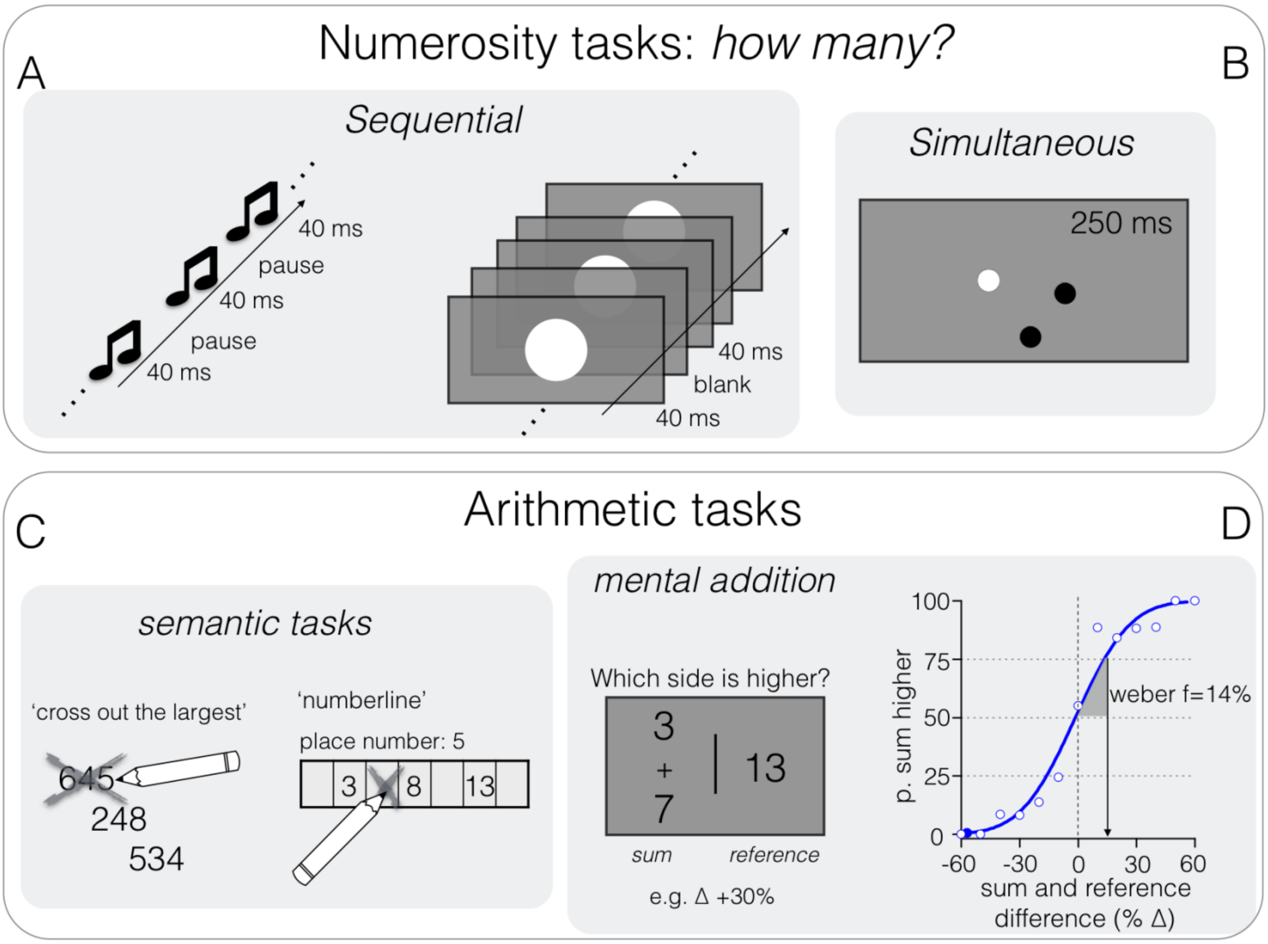



Figure 1. Tasks and paradigms. A and B) Each trial started with a fixation point (lasting until the experimenter pressed the space bar), followed by either a series of beeps or white disks flashes (A) or a cloud of dots simultaneously presented (B). Participants verbally reported perceived numerosity. C) Children were asked to solve a series of tasks where they had to recognize and cross the numerically larger digits among three, or to decide where a number should be placed in a sequence. D) Symbolic addition: on each trial, participants were asked to mentally add – as quickly as possible – the digits numbers on the left and compare the sum with that on the right (13 in this example), and indicate which side was numerically higher (right in the example). Weber Fraction (JND/PSE) measured precision: in the sample psychometric function reported, a WF of 0.14 indicates that the

sum of the two addenda had to be 14% higher or lower than reference to raise responses from chance (50%) to 75% correct responses. Stimuli remained until response.

## RESULTS

Task reliability. Table 1 reports split-half reliability levels (Pearson's r) for all the tasks for subject-by-subject analyses. Indexes were all reasonably good, ranging from 0.57 to 0.97. Regarding split-half subitizing reliability measures on pooled data, we found no statistically significant differences between average capacities calculated from the two data halves in both groups of children and adults (Table 2).

Table 1 – Split-half reliability indexes for children (C) and adults (A). Errors reflect standard errors.

| Tasks | Pearson's r |
|---|---|
| Dots | C: 0.64 ± 0.18<br>A: 0.65 ±0.19 |
| Flashes | C: 0.57 ±0.24<br>A: 0.64±0.20 |
| Tones | C: 0.68 ± 0.13<br>A: 0.74 ±0.17 |
| Mental | Precision (Weber Fraction) |

| | |
|---|---|
| addition | C: 0.58±0.18<br>A: 0.75±0.08<br><br>Speed (Reaction Times)<br>C: 0.97±0.006<br>A: 0.95±0.01 |

Table 2– Split-half averages and of subitizing capacities for children (C) and adults (A). Analyses were performed on pooled data. Errors reflect standard errors of the mean.

| | Subitizing capacities | | Average difference |
|---|---|---|---|
| Stimuli | First half | Second half | First-half Vs Second-half |
| Dots | C: 5.747± 0.223<br>A: 7.15±0.30 | C: 5.744± 0.226<br>A: 7.13±0.30 | C: 0.003±0.174 (p=0.98)<br>A: 0.012±0.29 (p=0.97) |
| Flashes | C: 4.53±0.18<br>A: 5.78±0.28 | C: 4.52±0.18<br>A: 5.79±0.28 | C: 0.011± 0.173 (p=0.93)<br>A: 0.006±0.24 (p=0.95) |
| Tones | C: 4.876±0.202<br>A: 6.24±0.29 | C: 4.876± 0.197<br>A: 6.22±0.29 | C: 0.0007± 0.177 (p=0.99)<br>A: 0.018±0.32 (p=0.96) |

**Goodness of fit**

The sigmoid fits describe well the data (Figure 1), with good coefficients of determination ($R^2$). Indeed, between participants average $R^2$ for children were 0.74 (SD 0.14, min 0.32), 0.73 (SD 0.17, min 0.33), and 0.76 (SD 0.37, min 0.37), for dots, flashes and tones estimations, respectively. $R^2$ fits for adults were on average 0.79 (SD 0.14, min 0.25), 0.76 (SD 0.17, min 0.42), and 0.75 (SD 0.18, min 0.38) for dots, flashes and tones estimations respectively. Some participants had at least one condition in which the $R^2$ was too poor to reliably estimate the subitizing limit. Similar Piazza et al's (2010), we adopted criterion of eliminating subjects with $R^2 < 0.25$. Ten children had a total of twelve low $R^2$ fit values (two fit for dots stimuli, six with flashes, four with sounds). Also two adults had poor $R^2$ (one for flash stimuli and one for sounds).

**Subitizing limits in children and adults**

We measured error rates (symbols on Figure 2) for estimating numerosity of dots arrays, sequences of flashes and auditory (tones) events. We fitted errors with sigmoid functions and took the inflection point as an index of subitizing limit. Figure 2 shows averaged results: all conditions clearly showed the classical subitizing signature, with low numbers characterized by lower error rates. This suggests that sequential events, like simultaneous spatial ensembles, can be subitized in early childhood.

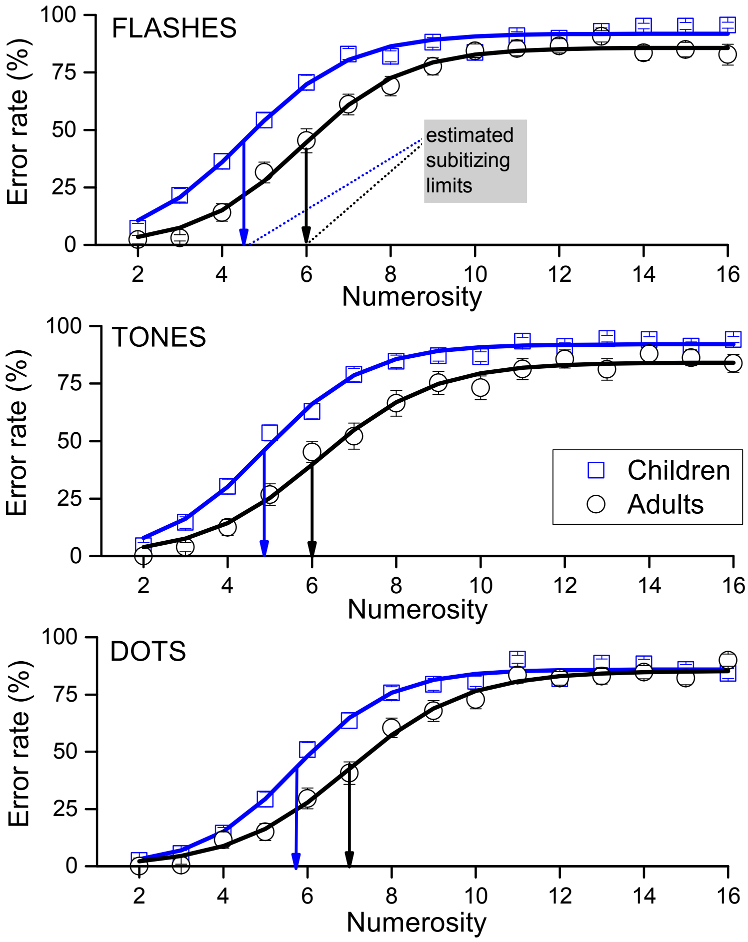



**Figure 2. Error rates as a function of numerosity.** Panels from top to bottom report performance for numerosity estimation of flashes, tones and arrays of dots, averaged across subjects. Data in blue refer to children, black to adults. Error bars are SEM. Lines are sigmoid functions. Arrows indicate subitizing capacities measured from the inflection point of the fitting functions.

The same fitting procedure was applied separately to each participant. Figures 3 A and B show the frequency distributions of subitizing limits across participants. On average, limits peaked around five/six items (overall averages, pooling together data for all stimuli, were 5.26 and 6.3 for children and adults respectively), a range often reported in the literature (for similar value in case of dots see Kaufman & Lord, 1949). More importantly, the distributions show much inter-subject variability. This replicates previous findings (Piazza et al., 2011), and suggests that the variance is large enough to run correlational analyses (described in the next paragraph).

In order to monitor developmental changes, we first computed average subitizing as a function of stimulus condition, separately for adults and children (Figure 3 C). From inspection, it is clear that the adult limits of subitizing were roughly one element higher than for primary school children. In order to statistically test the difference between children and adults we performed a 2 X 3 ANOVA (group: children, adults; stimuli: flashes, dots, sounds) with subitizing limits as dependent variable. The analyses confirmed that adults had higher subitizing limits ($F(1,397)=46$, $p<0.001$, $\eta 2=0.097$). The effect of stimuli was also significant ($F(2,397)=19$, $p<0.001$, $\eta 2=0.075$), with no interaction with group ($F(2,397)=0.085$, $p<0.91$, $\eta 2=0.000$), suggesting that some subitizing measures differ from others, and the difference was constant across the group. For both children and adults, simultaneous subitizing limits were higher than those for temporal stimuli, while visual and auditory temporal limits were very similar to each other (Table 3).

**Table 3– Difference between averaged subitizing capacities across stimuli for children (C) and adults (A)**

| Stimuli | Difference | 95% CI of the difference Low | High |
|---|---|---|---|
| Dots vs Flashes | C: 1.16***<br>A: 1.09** | C: 0.69<br>A: 0.16 | C: 1.63<br>A: 2.03 |
| Dots vs Tones | C: 0.87***<br>A: 0.97* | C: 0.41<br>A: 0.04 | C: 1.34<br>A: 1.9 |
| Tones vs Flashes | C: 0.28 n.s.<br>A: 0.13 n.s. | C: -0.75<br>A: -0.8 | C: 0.17<br>A: 1.06 |

Two tail t-test p-values = *p<0.05 ; ** p<0.01 ;  ***p<0.001

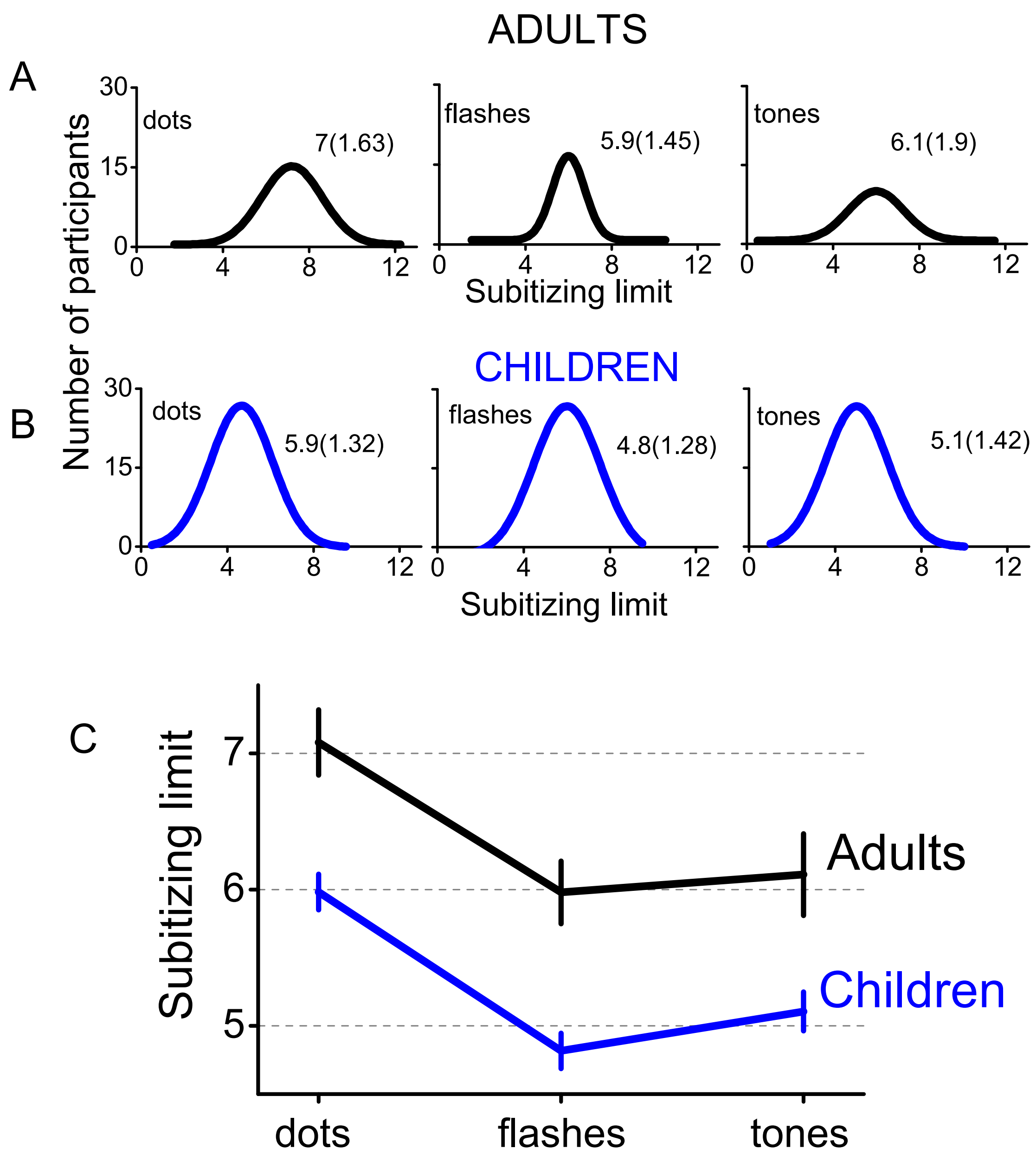



**Figure 3. Frequency distributions of subitizing capacities.** A and B: Panels from left to right report data for subitizing of different stimuli: simultaneous numerosity, sequential numerosity for visual and auditory stimuli respectively. The first row shows data for adults (in black) whilst children data are shown in the second row (in blue). C) Average subitizing

capacity as a function of stimuli, for children (blue) and adults (black). Error bars show SEM.

We further tested developmental trajectories of simultaneous and sequential subitizing, correlating age and subitizing limits. The results confirmed that all subitizing limits significantly increase from childhood to adulthood (Figure 4 A, B and C, black regression lines). We then looked at developmental changes within the two groups, separately. Within the child sample, only subitizing limits for visual sequential stimuli clearly improved with age, with auditory subitizing approaching the significance level and no significant correlation for subitizing of simultaneous numerosity (Pearson zero-order correlations, one-tail p-values; dots: $r=0.04$, $p=0.34$, $LBF=-1$; flashes: $r=0.29$, $p=0.002$, $LBF=0.8$; tones: $r=0.14$, $p=0.09$, $LBF=-0.23$; Figure 3 D,E and F, red regression lines). For adult participants, no condition correlates with age (all $p>0.05$). These results suggest that at around 7 years old, all subitizing limits except those for sequential numerosity have fully matured. These additional results highlight differences between simultaneous and sequential subitizing in both developmental trajectories and system capacity.

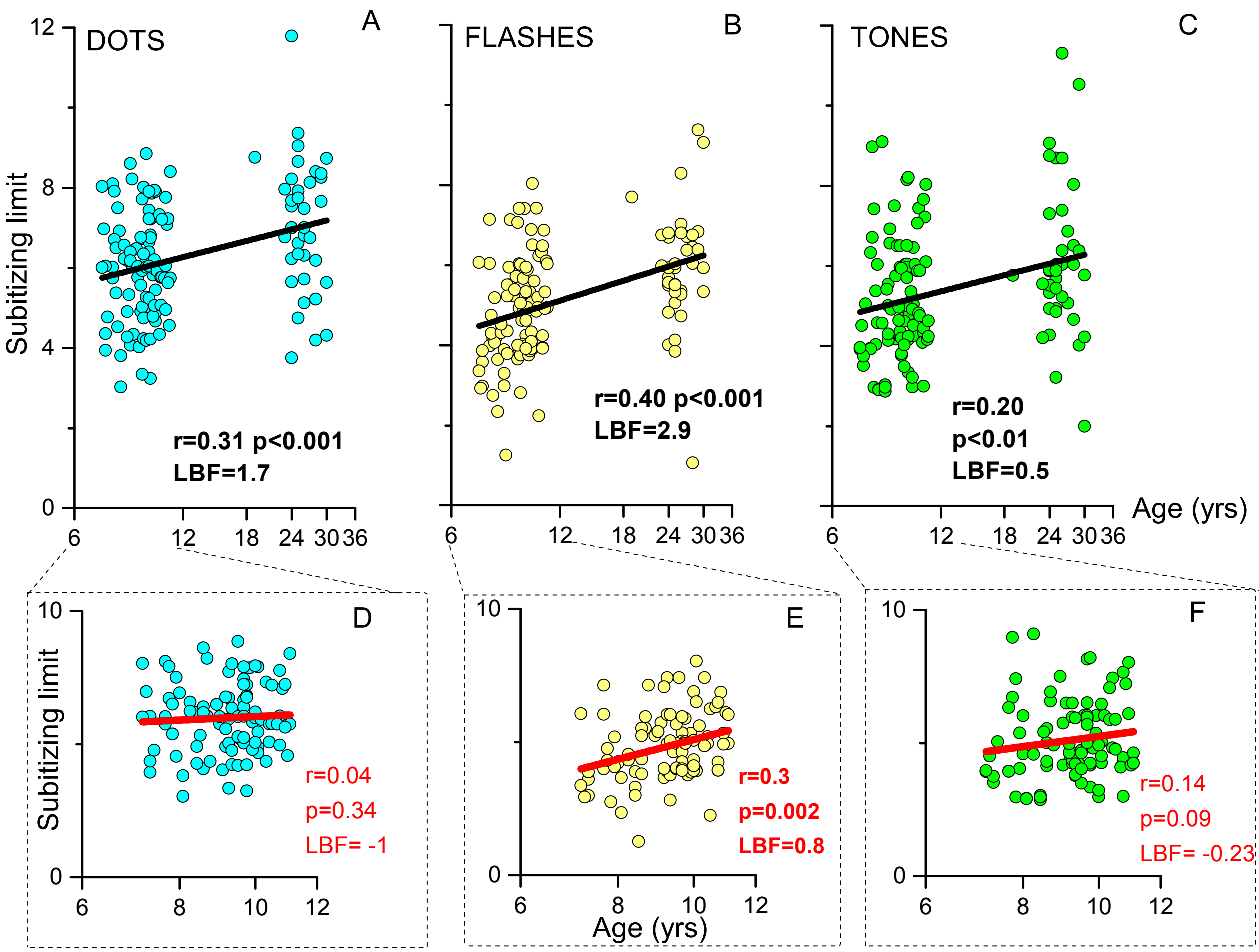



**Figure 4. Developmental trajectories.** Panels from left to right report data for subitizing capacities as a function participant's age for children and adults.

## Correlation between simultaneous and sequential subitizing

The results so far show that both adults and children can subitize simultaneous and sequential stimuli. We also found that simultaneous subitizing outperforms sequential subitizing, regardless of the sensory modality of the stimuli, and that subitizing capacities for different kind of stimuli develop differently during childhood, suggesting different systems. Here we investigated this possibility further by correlating simultaneous and sequential subitizing limits. The results show that child subitizing limits for sequential stimuli positively correlate between each other: children with higher subitizing limits for sequential sequences of visual events also have higher limits for sequences of tones (r=0.28, p= 0.003, LBF=0.54, Table 4). However, subitizing for simultaneous stimuli did not correlate with any of the sequential conditions (dots vs flashes r=0.07, p=0.23, LBF = −1; dots vs tones r=0.21, p=0.02, LBF = −0.19). The positive significant correlation between sequential stimuli was robust, as it remains significant even when the effect of age and non-verbal IQ were simultaneously controlled for (rp=0.28, p= 0.006, Table 4, below diagonal). No significant correlation was found for adult participants (dots vs flashes r= 0.28 p=0.08, LBF = −0.28; dots vs tones r=0.27, p=0.09, LBF = −0.92; flashes vs tones r=0.07, p=0.64, LBF= −0.89). These analyses are in line with those previously reported in this manuscript to support the idea of two different systems for simultaneous and sequential subitizing.

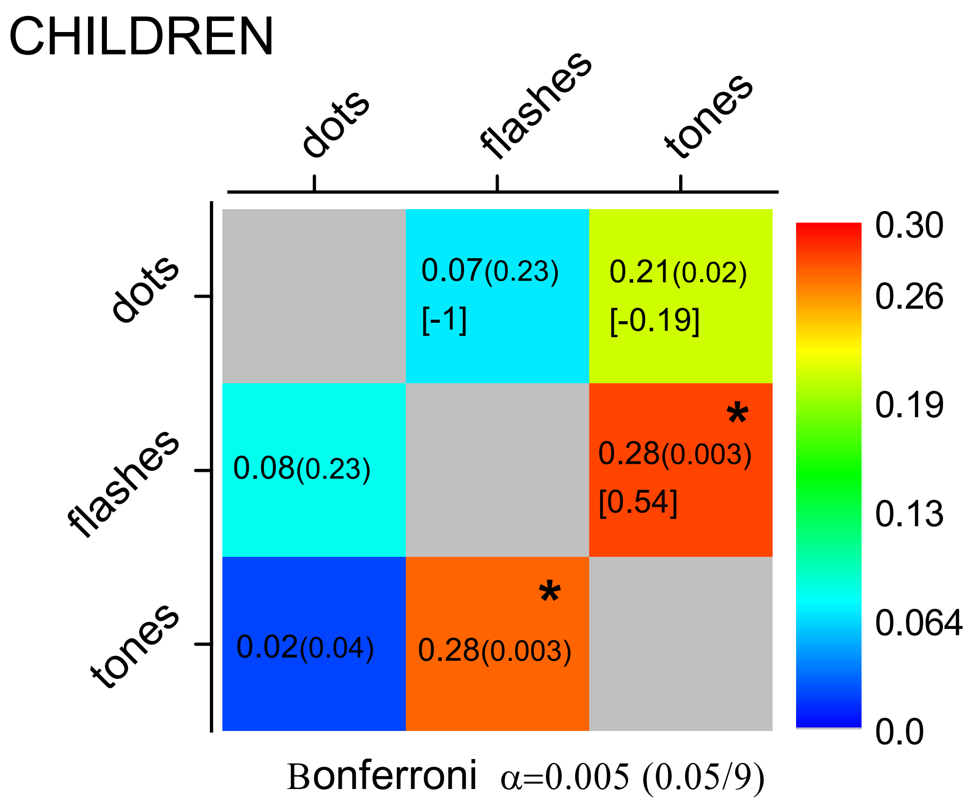
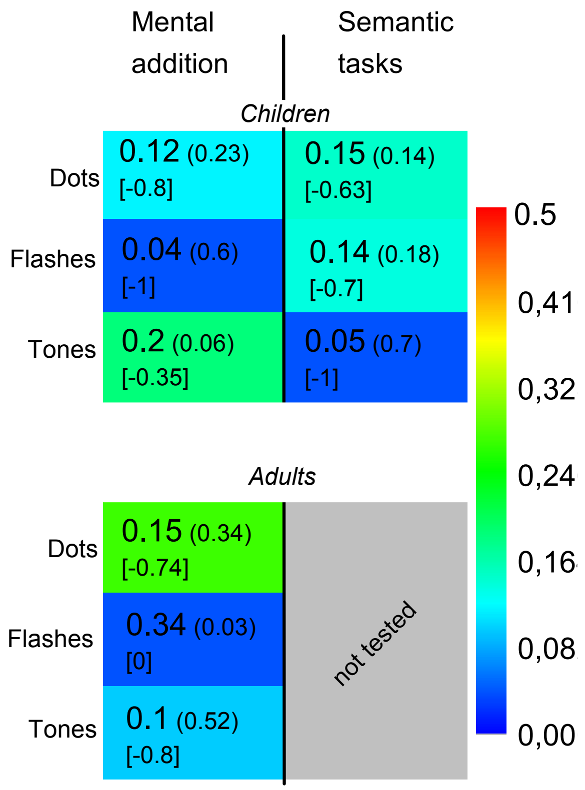


**Table 4. Correlations between subitizing limits.** Above diagonal: Zero-order Pearson correlations. Below diagonal: Pearson partial correlations (Age and Raven controlled). One-tail p-values are reported in round brackets. Log Bayes Factors (LBF) are reported in square brackets.

**Correlation between subitizing limits and mathematical abilities**

We then went on to investigate whether subitizing limits correlate with mathematical skills. At a first level of analyses, we correlated children and adult mathematical scores with subitizing limits. The results show that none of the subitizing measures correlate with any of the mathematical scores (overall indexes for mental addition and semantic tasks), neither in children nor in adults (Table 5). Beside the high p-values, Bayes Factors clearly speak in favour of the null hypothesis, with values less than −0.5 considered substantial evidence in favour of the null hypothesis, less than −1 strong evidence. It is important to note that we have recently shown the very same math scores correlated with spatial simultaneous numerosity estimation and discrimination precision levels for numerosities higher than the subitizing range (see figure 6 in Anobile et al., 2017). This clearly shows that the lack of correlation observed here does not depend on problems with measuring mathematical abilities.



| | Mental addition | Semantic tasks |
|---|---|---|
| *Children* | | |
| Dots | 0.12 (0.23) [-0.8] | 0.15 (0.14) [-0.63] |
| Flashes | 0.04 (0.6) [-1] | 0.14 (0.18) [-0.7] |
| Tones | 0.2 (0.06) [-0.35] | 0.05 (0.7) [-1] |
| *Adults* | | |
| Dots | 0.15 (0.34) [-0.74] | not tested |
| Flashes | 0.34 (0.03) [0] | not tested |
| Tones | 0.1 (0.52) [-0.8] | not tested |



**Table 5. Zero-order Pearson correlations between subitizing capacities and math abilities in children and adults.** Two-tail p-values are reported in round brackets. Adult Bonferroni corrected α=0.008 (0.05/6); Child Bonferroni corrected α =0.005 (0.05/9). Log Bayes Factors (LBF) are reported in square brackets.

We then investigated whether a significant correlation may have emerged by considering response speed or accuracy in the math tasks independently. We thus performed again the correlation analysis (two-tailed, zero-order Pearson) between subitizing limits and the mental addition task (the only one shared between children and adults), considering separately speed (raw values of RTs) and precision level (raw values of WFs). The results are reported in Table 5. It is clear that no significant correlations between subitizing and mental addition proficiency emerged in any dimension or group of subjects.

**Table 5. Zero-order Pearson correlations between subitizing capacities and mental addition proficiency in children and adults.**

| | dots | flashes | tones |
|---|---|---|---|
| | | CHILDREN | |
| Speed (secs) | 0.06 (0.5) [−1.8] | −0.13 (0.17) [−1.3] | −0.1 (0.29) [−1.5] |
| Precision (Wfs) | −0.13 (0.21) [−1.3] | −0.12 (0.22) [−1.16] | −0.17 (0.08) [−0.55] |
| | | ADULTS | |
| Speed (secs) | −0.07 (0.66) [−1.5] | −0.16 (0.32) [−1.13] | −0.19 (0.23) [−0.9] |
| Precision (Wfs) | −0.16 (0.33) [−1.1] | −0.34 (0.03) [0.5] | 0.04 (0.81) [−1.5] |

Two-tailed p-values are reported in round brackets; Log Bayes Factors (LBF) are reported in square brackets. Bonferroni corrected α=0.008 (0.05/6).

Correlational analyses may have missed potential clumping in the data and thus obscured the presence of potential distinct subpopulations in the sample. To explore this possibility, we consider child data (the larger sample size) and ran a two-step cluster analysis by considering all the available variables: age, IQ, mental calculation and math score in the paper and pencil task. The analysis identified three clusters containing 28, 30 and 38 participants. To check whether subitizing limits differ among these three sub-populations, we ran a 3 (clusters = 1, 2 and 3) X 3 (subitizing capacity = dots, flashes, tones) repeated measures ANOVA. The interaction between factors was found to be not significant ($F(4,283) = 0.54$, $p = 0.7$), suggesting that these groups did not perform differently in any of the three different subitizing tasks. We then tested for significant differences in all possible factor combinations with a series of post-hoc t-tests. Also these analyses confirmed that none of the subitizing measures differed across participant clusters (min $p = 0.06$, Bonferroni alpha corrected= 0.005).

We then applied a more extreme approach: we split the child sample into two sub-samples, collapsing higher and lower math skilled subjects. Because the performance on two symbolic math tasks correlated well between each other ($r=0.43$, $p<0.001$), we built a summary math index by averaging the z-scores of the two tests, and used this index value to compute child math percentiles. Those above the 85th percentile were assigned to the "high-math" group (N14, average z-score 0.76), those below the 15th assigned to the "low-math" group (N14, average z-score −0.79). With a non-parametric sample-with-replacement bootstrap technique (10000 iterations) we built the two-math distributions shown in Figure 5: on each iteration math z-scores were resampled (with replacement) separately for each math group, and the average z-scores computed. At the same time we calculated subitizing limits, separately for each stimulus condition and math group. Figure 5 (C, D and E) shows the average subitizing frequency distributions for the two math groups (red "low", green "high"). Those distributions largely overlap with virtually no subitizing advantage for high-math children for any of the stimulus conditions. We statistically computed the difference between those distributions, counting the times that, on each of the 10000 iterations, the difference between the averaged subitizing capacities were higher than zero (one-tail p-value). All p-values (reported in Figure 5 C, D, E) were near 0.5, showing clearly that they were not statistically different, robustly reinforcing our finding of no correlation between subitizing and math capacities.

We repeated the analysis with even more conservative criteria, considering only the 5% tails of the math distribution (5-95 percentile, samples size was $N=5$ and $N=6$ for $\leq$ 5% and $\geq$ 95% respectively) or the 10% (10-90 percentile, $N=11$ each group). Even in these cases none of the subitizing measures differed between groups (all $p\approx0.5$). We then checked whether an automatized analysis might have identified the same group of subjects than our custom procedure by running a cluster analysis based on math abilities (overall index). The obtained dendrogram revealed a maximum distance between a small group of participants (with particularly low scores on math ability) and the rest of the participants. Interestingly, this group consisted of the very same participants included in the 5$^{th}$ percentile of math proficiency distribution (that as shown above did show a subitizing range rather identical to all other participants, even to those in the 5$^{th}$ percentile of the highest match scores.

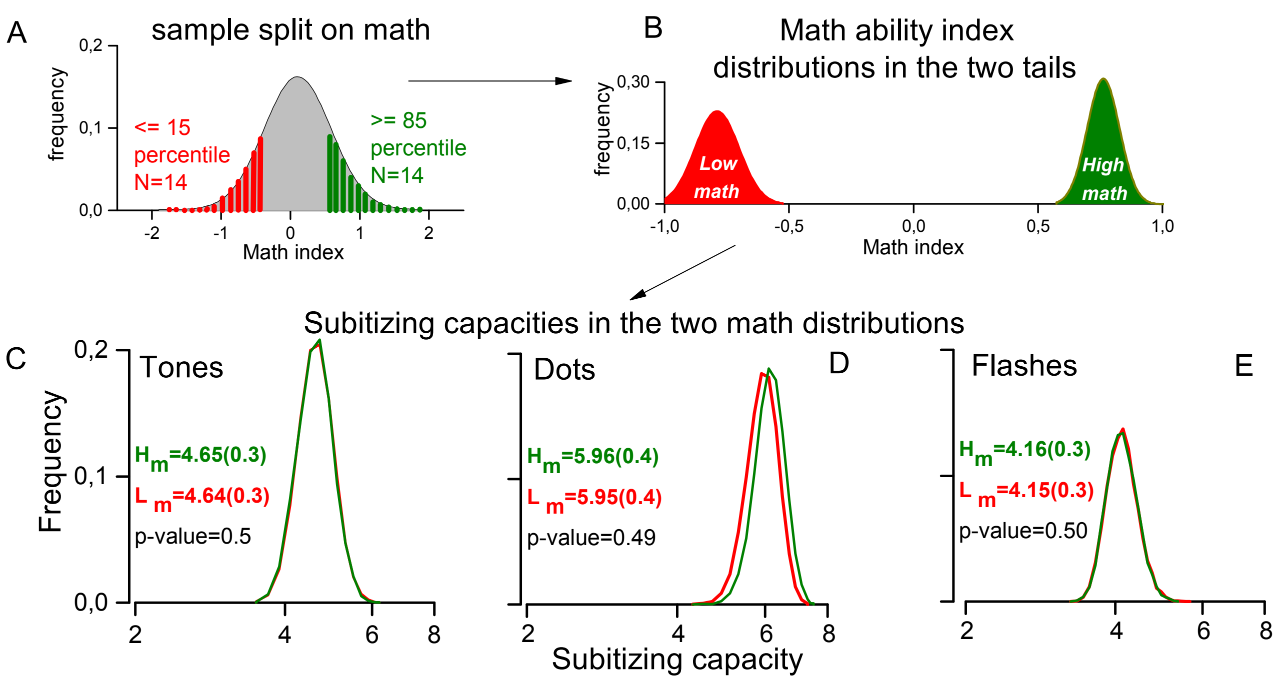



**Figure 5. Subitizing limits for "low" and "high" math children.** A) Frequency distribution of child math scores, with tails (14 each) falling in the selected percentile ranges: 1) ≥ 85, green, "high math" and 2) ≤15, red, "low math". B) Child math scores frequency distributions inside the two percentile samples. C, D and E) Subitizing limit distributions separated for the two math samples (Hm = high math, green; Lm = low math, red). Values reflect: average and SD (in brackets) subitizing capacities and associated one tail p-values.

## DISCUSSION

We measured subitizing limits (as the point of discontinuity in estimation error rates) for both simultaneous and sequential numerosity for primary school children and adults. We also measured formal math capacity (mental addition and digit magnitude manipulation) and non-verbal reasoning abilities (Raven matrices). We found that: 1) regardless of stimuli sensory modality and presentation format (simultaneous or temporal), adults subitize one item more than children; 2) subitizing limits for simultaneous stimuli do not change between 7 and 11 years, while subitizing limits for sequential stimuli (particularly in case of visual flashes) significantly improve in this period; 3) simultaneous and sequential subitizing limits do not correlate with each other, but sequential subitizing for visual and auditory stimuli do correlate, even over the age range 7-11; 4) in neither group of subjects did any form of subitizing measure correlate with math skills.

### Subitizing and math abilities

The null correlation between subitizing and math scores may sound counterintuitive, especially for simultaneous subitizing (dots), but some issues are worth considering. First, the literature linking symbolic math and subitizing limits is not very solid. The most robust studies describing a link between math abilities and subitizing encompass patients with math impairments associated with neurological disorders such as cerebral palsy (Arp & Fagard, 2005; Arp et al., 2006), Turner's syndrome (Bruandet, Molko, Cohen, & Dehaene, 2004), Williams and Down syndrome (Paterson, Girelli, Butterworth, & Karmiloff-Smith, 2006), Gerstmann's syndrome (Cipolotti, Butterworth, & Denes, 1991; Lemer, Dehaene, Spelke, & Cohen, 2003). Although they are informative, these studies suffer from the caveat that non-numerical deficits associated with those neurological disorders may have impacted subitizing.

Another important point to consider is that we discouraged serial counting by using a very fast presentation time (for dots) or high temporal rate (for flashes and sounds), and did not require speeded responses. However, our error rate was very similar to that commonly documented in the literature, and measures of reliability were always high. In the light of these data we can reasonably assume that the methods we applied to measure subitizing are sensitive and robust. We quantified the likelihood of this null correlation by means of Bayes factor. Log Bayes factors for correlations between subitizing limits and math skill were all clearly negative and mainly near −1 (Table 5, square brackets), indicating strong evidence in favour of the null hypothesis of zero correlation (Wetzels & Wagenmakers, 2012). It is also worth noting that object serial counting speed has been found to be a good and stable marker of dyscalculia (Gray & Reeve, 2014; Reeve, Reynolds, Humberstone, & Butterworth, 2012), leaving open the interesting possibility that leaving open the interesting possibility that the link between math and subitizing occurs only when counting is used.

Another important point is the heterogeneity of the indexes used to measure subitizing efficiency as well as math skills. All measures have advantages and disadvantages, so the choice of measure is dictated by the experimental goals of the studies. For example, to measure subitizing proficiency, some studies have used the RT acceleration as a function of numerosity (linear fit slope) in the small number range. This method has the assumption that an increase of RTs reflects a less efficient subitizing system. With this index Schleifer and Landerl (2011) found that dyscalculic children had higher slope than controls, suggesting these subjects had an impaired subitizing system and used inefficient serial counting even in this small number range. Nevertheless, at the error rates reported in their study (see Figure 2), it is evident that even dyscalculics showed a marked subitizing effect and that subitizing limits (the point of discontinuity) were not evidently different from controls. Here, as in other studies (Green & Bavelier, 2003; Piazza et al., 2011), we focused on this latter parameter (point of discontinuity).

In this study we focused on arithmetic tasks that particularly tap into rote memorization and semantic skills, abilities that have been previously found to correlate with precision in estimating numerosity above the subitizing range. However, we are aware that these results do not exclude the possibility that other arithmetic tasks might instead correlate with subjects' ability to subitize. As mentioned above, no causal studies investigated the link between math and subitizing. Curiously the only cognitive capacity that was causally related to subitizing was not numerical. In their seminal paper Green and Bavelier (2003) showed that boosting visual attention capacities by playing action video games enlarged the subitizing capacity of adults. In line with that, depriving attention greatly degrades subitizing, far more than estimation (Anobile et al., 2012; Burr, Anobile, & Turi, 2011; Burr et al., 2010; Pagano, Lombardi, & Mazza, 2014; Railo, Koivisto, Revonsuo, & Hannula, 2008; Vetter et al., 2008). Other studies point to a crucial role of non-numerical factors such as visual working memory (Piazza et al., 2011) and stimulus spatial configural processing (Ashkenazi et al., 2013; Krajcsi, Szabo, & Morocz, 2013; Mandler & Shebo, 1982).

For the aims of the present study we reanalysed recent data collected for other purposes. In a previous study we only considered data in the estimation range, carefully avoiding subitizing (Anobile et al., 2017), and detected a good correlation between children simultaneous (spatial arrays) numerosity estimation precision and math. It is important to note that formal math tasks and scores in the two studies were identical, and therefore the lack of correlation found here cannot be accounted for by difficulties in measurement of arithmetic abilities. In any case, since mathematics is not a single concept, the possibility still remains open that certain skills may be exclusively linked to the ANS and others to the subitizing system. In previous reports, our group has proposed the ANS and subitizing overlap, but  for very low numerosities (subitizing range) precision is boosted by attentional resources making it particularly fast and precise. Depriving attentional resources, precision level in the subitizing range approaches that of the estimation range (Anobile et al., 2012; Burr et al., 2010). Moreover, while in normal conditions subitizing is not susceptible to numerosity adaptation, it is adaptable when attention is deprived in a dual-task paradigm (Burr et al., 2011). It would be interesting to test whether performance within the subitizing range, measured under dual task conditions, correlated with math abilities.

## Subitizing across development

The results also show that adult subitizing limits were constantly one item larger than those of children across all the stimuli format conditions (spatial arrays or temporal sequences). Larger subitizing may arise from genuine development of the subitizing system(s), but could also arise from more efficient domain-general mechanisms related to the subitizing phenomena (i.e. attentional and/or WM capacities). It would be interesting to test whether the detected developmental differences hold even after regressing out domain general non-numerical abilities.

## Subitizing across sensory modalities and stimuli format

The present results confirm previous studies showing that adults can subitize auditory and visual sequential stimuli (Camos & Tillmann, 2008; Repp, 2007) and go on to show that this ability is present in primary school children. In children, the capacity to subitize audio and visual sequential events are positively correlated with each other, indicating a common system for perception of sequential stimuli. This correlation was strong, with LBF near 0.5, robust enough to survive Bonferroni correction, and remained significant even when the important covariates of age and non-verbal reasoning scores (Raven matrices) were controlled for. With adults we found no significant correlations between subitizing capacities. In light of these results we might hypothesize that sequential and simultaneous subitizing are subserved by separate mechanisms. As auditory and visual sequential subitizing capacities are linked in children but not in adults, this may suggest that the “sequential subitizing system” starts as a cross-sensory system that differentiates later on. Since in our sample of children we did not find a correlation between simultaneous and sequential subitizing, this hypothesis predicts that in younger children this correlation should exist. To indicate which factors may cause the hypothesised differentiation is difficult to say, but we may speculate that it should reflect a gradually reduced cross-talk between general domain skills across different sensory modalities (for example auditory and visual attentional and/or working memory resources). It would be interesting to devise future studies to test all these hypotheses. Moreover, a lack of correlation cannot be interpreted as definitive prove of separate mechanisms. Other studies, perhaps using causative methods, are needed to demonstrate separate mechanisms for simultaneous and sequential subitizing, as well as to define their relationship with math. As a complimentary way to test for our current results, it would be interesting to look for neuropsychological dissociations in patients with brain lesions as the represented data clearly predict the possibility of deficits selectively affecting simultaneous or sequential subitizing abilities.

## Conclusions

Overall these results suggest that although enumeration of simultaneous (dots) and sequential (sounds and flashes) stimuli both shows the classical subitizing performance advantage, they may be sub-served by separate systems. Furthermore, subitizing limits for dots, flashes as well as sounds does not seem to be related to numerical abilities (at least with those measured in the present study).

## References


Agrillo, C., Miletto Petrazzini, M. E., & Bisazza, A. (2017). Numerical abilities in fish: A methodological review. *Behav Processes*. doi:10.1016/j.beproc.2017.02.001

Anobile, G., Arrighi, R., Castaldi, E., Grassi, E., Pedonese, L., PA, M. M., & Burr, D. C. (2017). Spatial but Not Temporal Numerosity Thresholds Correlate With Formal Math Skills in Children. *Dev Psychol*. doi:10.1037/dev0000448

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

Anobile, G., Stievano, P., & Burr, D. C. (2013). Visual sustained attention and numerosity sensitivity correlate with math achievement in children. *J Exp Child Psychol, 116*(2), 380-391. doi:10.1016/j.jecp.2013.06.006

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 Res, 74*, 102-109. doi:10.1016/j.visres.2012.06.005

Arp, S., & Fagard, J. (2005). What impairs subitizing in cerebral palsied children? *Developmental Psychobiology, 47*(1), 89-102. doi:10.1002/dev.20069

Arp, S., Taranne, P., & Fagard, J. (2006). Global perception of small numerosities (subitizing) in cerebral-palsied children. *J Clin Exp Neuropsychol, 28*(3), 405-419. doi:10.1080/13803390590935426

Arrighi, R., Lunardi, R., & Burr, D. (2011). Vision and audition do not share attentional resources in sustained tasks. *Front Psychol, 2*, 56. doi:10.3389/fpsyg.2011.00056

Arrighi, R., Togoli, I., & Burr, D. C. (2014). A generalized sense of number. *Proc Biol Sci, 281*(1797). doi:10.1098/rspb.2014.1791

Ashkenazi, S., & Henik, A. (2012). Does attentional training improve numerical processing in developmental dyscalculia? *Neuropsychology, 26*(1), 45-56. doi:10.1037/a0026209

Ashkenazi, S., Mark-Zigdon, N., & Henik, A. (2013). Do subitizing deficits in developmental dyscalculia involve pattern recognition weakness? *Developmental Science, 16*(1), 35-46. doi:10.1111/j.1467-7687.2012.01190.x

Askenazi, S., & Henik, A. (2010). Attentional networks in developmental dyscalculia. *Behav Brain Funct, 6*, 2. doi:10.1186/1744-9081-6-2

Biancardi, A., Bachmann, C., & Nicoletti, C. ( 2016). *Batteria per la discalculia evolutiva (BDE2)*. Trento: centro studi erickson.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis, 10*(4), 433-436.

Bruandet, M., Molko, N., Cohen, L., & Dehaene, S. (2004). A cognitive characterization of dyscalculia in Turner syndrome. *Neuropsychologia, 42*(3), 288-298.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: inhibition, switching, and working memory. *Dev Neuropsychol, 19*(3), 273-293. doi:10.1207/S15326942DN1903_3

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

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *J Vis, 10*(6), 20. doi:10.1167/10.6.20

Camos, V., & Tillmann, B. (2008). Discontinuity in the enumeration of sequentially presented auditory and visual stimuli. *Cognition, 107*(3), 1135-1143. doi:10.1016/j.cognition.2007.11.002

Carey, S. (2002). Cognitive foundations of arithmetic: evolution and ontogenesis. *Mind and Language, 16*, 37–55 doi:10.1111/1468-0017.00155

Ceulemans, A., Titeca, D., Loeys, T., Hoppenbrouwers, K., Rousseau, S., & Desoete, A. (2014). Enumeration of small and large numerosities in adolescents with mathematical learning disorders. *Res Dev Disabil, 35*(1), 27-35. doi:10.1016/j.ridd.2013.10.018

Cicchini, G. M., Anobile, G., & Burr, D. C. (2016). Spontaneous perception of numerosity in humans. *Nat Commun, 7*. doi:10.1038/ncomms12536

Cipolotti, L., Butterworth, B., & Denes, G. (1991). A specific deficit for numbers in a case of dense acalculia. *Brain, 114 ( Pt 6)*, 2619-2637.

Cowan, N., Morey, C. C., AuBuchon, A. M., Zwilling, C. E., & Gilchrist, A. L. (2010). Seven-year-olds allocate attention like adults unless working memory is overloaded. *Dev Sci, 13*(1), 120-133. doi:10.1111/j.1467-7687.2009.00864.x

Dacke, M., & Srinivasan, M. V. (2008). Evidence for counting in insects. *Anim Cogn, 11*(4), 683-689. doi:10.1007/s10071-008-0159-y

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquiere, P. (2009). Working memory and individual differences in mathematics achievement: a longitudinal study from first grade to second grade. *J Exp Child Psychol, 103*(2), 186-201. doi:10.1016/j.jecp.2009.01.004

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

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cogn Neuropsychol, 20*(3), 487-506. doi:10.1080/02643290244000239

Desoete, A., & Grégoire, J. (2006). Numerical competence in young children and in children with mathematics learning disabilities
. *Learning and Individual Differences, 16*, 351-367. doi:10.1016/j.lindif.2006.12.006

Ditz, H. M., & Nieder, A. (2015). Neurons selective to the number of visual items in the corvid songbird endbrain. *Proc Natl Acad Sci U S A, 112*(25), 7827-7832. doi:10.1073/pnas.1504245112

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

Ferrand, L., Riggs, K. J., & Castronovo, J. (2010). Subitizing in congenitally blind adults. *Psychon Bull Rev, 17*(6), 840-845. doi:10.3758/PBR.17.6.840

Fischer, B., Gebhardt, C., & Hartnegg, K. (2008). Subitizing and visual counting in children with problems acquiring basic arithmetic skills . *Optometry and Vision Development, 39*, 24-29.

Gallace, A., Tan, H. Z., Haggard, P., & Spence, C. (2008). Short term memory for tactile stimuli. *Brain Res, 1190*, 132-142. doi:10.1016/j.brainres.2007.11.014

Gray, S. A., & Reeve, R. A. (2014). Preschoolers' Dot Enumeration Abilities Are Markers of Their Arithmetic Competence. *PLoS One, 9*(4). doi:ARTN e94428 10.1371/journal.pone.0094428

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*(6939), 534-537. doi:10.1038/nature01647

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proc Natl Acad Sci U S A, 106*(25), 10382-10385. doi:10.1073/pnas.0812142106

Kaufman, E. L., & Lord, M. W. (1949). The discrimination of visual number. *Am J Psychol, 62*(4), 498-525.

Krajcsi, A., Szabo, E., & Morocz, I. A. (2013). Subitizing Is Sensitive to the Arrangement of Objects. *Experimental Psychology, 60*(4), 227-234. doi:10.1027/1618-3169/a000191

Lemer, C., Dehaene, S., Spelke, E., & Cohen, L. (2003). Approximate quantities and exact number words: dissociable systems. *Neuropsychologia, 41*(14), 1942-1958.

Mandler, G., & Shebo, B. J. (1982). Subitizing - an Analysis of Its Component Processes. *Journal of Experimental Psychology-General, 111*(1), 1-22. doi:Doi 10.1037/0096-3445.111.1.1

McLachlan, N. M., Marco, D. J. T., & Wilson, S. J. (2012). Pitch Enumeration: Failure to Subitize in Audition. *PLoS One, 7*(4). doi:ARTN e33661
10.1371/journal.pone.0033661

Melcher, D., & Piazza, M. (2011). The role of attentional priority and saliency in determining capacity limits in enumeration and visual working memory. *PLoS One, 6*(12), e29296. doi:10.1371/journal.pone.0029296

Nieder, A. (2016). The neuronal code for number. *Nat Rev Neurosci, 17*(6), 366-382. doi:10.1038/nrn.2016.40

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

Pagano, S., Lombardi, L., & Mazza, V. (2014). Brain dynamics of attention and working memory engagement in subitizing. *Brain Res, 1543*, 244-252. doi:10.1016/j.brainres.2013.11.025

Passolunghi, M. C., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. *J Exp Child Psychol, 88*(4), 348-367. doi:10.1016/j.jecp.2004.04.002

Paterson, S. J., Girelli, L., Butterworth, B., & Karmiloff-Smith, A. (2006). Are numerical impairments syndrome specific? Evidence from Williams syndrome and Down's syndrome. *J Child Psychol Psychiatry, 47*(2), 190-204. doi:10.1111/j.1469-7610.2005.01460.x

Petrazzini, M. E. M., Agrillo, C., Izard, V., & Bisazza, A. (2016). Do Humans (Homo sapiens) and Fish (Pterophyllum scalare) Make Similar Numerosity Judgments? *Journal of Comparative Psychology, 130*(4), 380-390. doi:10.1037/com0000045

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences, 14*(12), 542-551. doi:Doi 10.1016/J.Tics.2010.09.008

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

Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition, 121*(1), 147-153. doi:10.1016/j.cognition.2011.05.007

Plaisier, M. A., Bergmann Tiest, W. M., & Kappers, A. M. (2009). One, two, three, many - subitizing in active touch. *Acta Psychol (Amst), 131*(2), 163-170. doi:10.1016/j.actpsy.2009.04.003

Plaisier, M. A., & Smeets, J. B. (2011). Haptic subitizing across the fingers. *Atten Percept Psychophys, 73*(5), 1579-1585. doi:10.3758/s13414-011-0124-8

Plaisier, M. A., Tiest, W. M. B., & Kappers, A. M. L. (2010). Grabbing subitizing with both hands: bimanual number processing. *Experimental Brain Research, 202*(2), 507-512. doi:10.1007/s00221-009-2146-1

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spat Vis, 3*(3), 179-197.

Raghubara, K. P., Barnesb, M. A., & Hechtb, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*(2), 110-122.

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

Reeve, R., Reynolds, F., Humberstone, J., & Butterworth, B. (2012). Stability and change in markers of core numerical competencies. *J Exp Psychol Gen, 141*(4), 649-666. doi:10.1037/a0027520

Repp, B. H. (2007). Perceiving the numerosity of rapidly occurring auditory events in metrical and nonmetrical contexts. *Percept Psychophys, 69*(4), 529-543.

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science, 19*(6), 607-614. doi:Doi 10.1111/J.1467-9280.2008.02130.X

Riggs, K. J., Ferrand, L., Lancelin, D., Fryziel, L., Dumur, G., & Simpson, A. (2006). Subitizing in tactile perception. *Psychological Science, 17*(4), 271-272. doi:10.1111/j.1467-9280.2006.01696.x

Rugani, R., Vallortigara, G., Priftis, K., & Regolin, L. (2015). Animal cognition. Number-space mapping in the newborn chick resembles humans' mental number line. *Science, 347*(6221), 534-536. doi:10.1126/science.aaa1379

Schleifer, P., & Landerl, K. (2011). Subitizing and counting in typical and atypical development. *Dev Sci, 14*(2), 280-291.

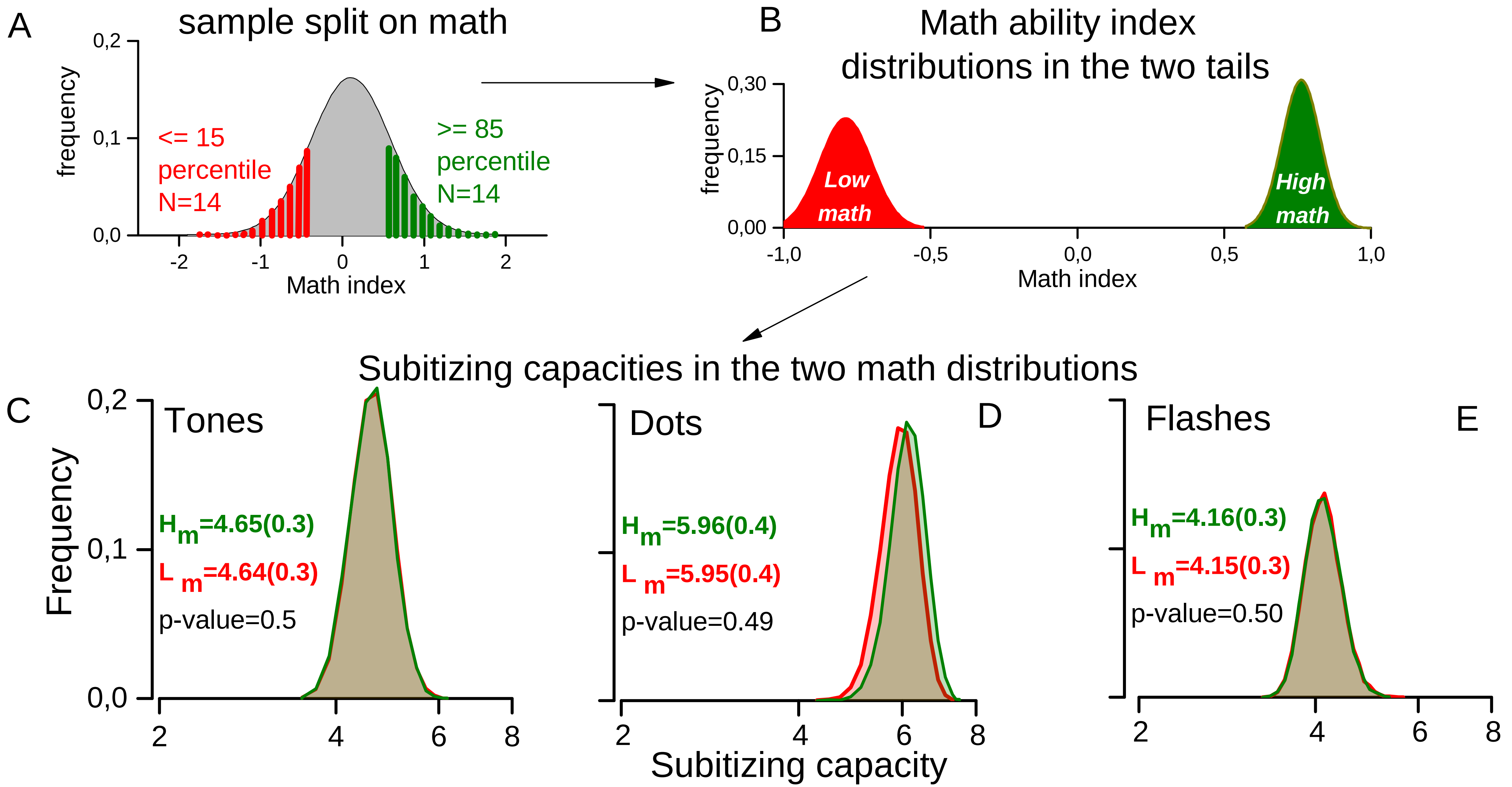
Steele, A., Karmiloff-Smith, A., Cornish, K., & Scerif, G. (2012). The multiple subfunctions of attention: differential developmental gateways to literacy and numeracy. *Child Dev, 83*(6), 2028-2041. doi:10.1111/j.1467-8624.2012.01809.x

Toll, S. W., Kroesbergen, E. H., & Van Luit, J. E. (2016). Visual working memory and number sense: Testing the double deficit hypothesis in mathematics. *Br J Educ Psychol, 86*(3), 429-445. doi:10.1111/bjep.12116

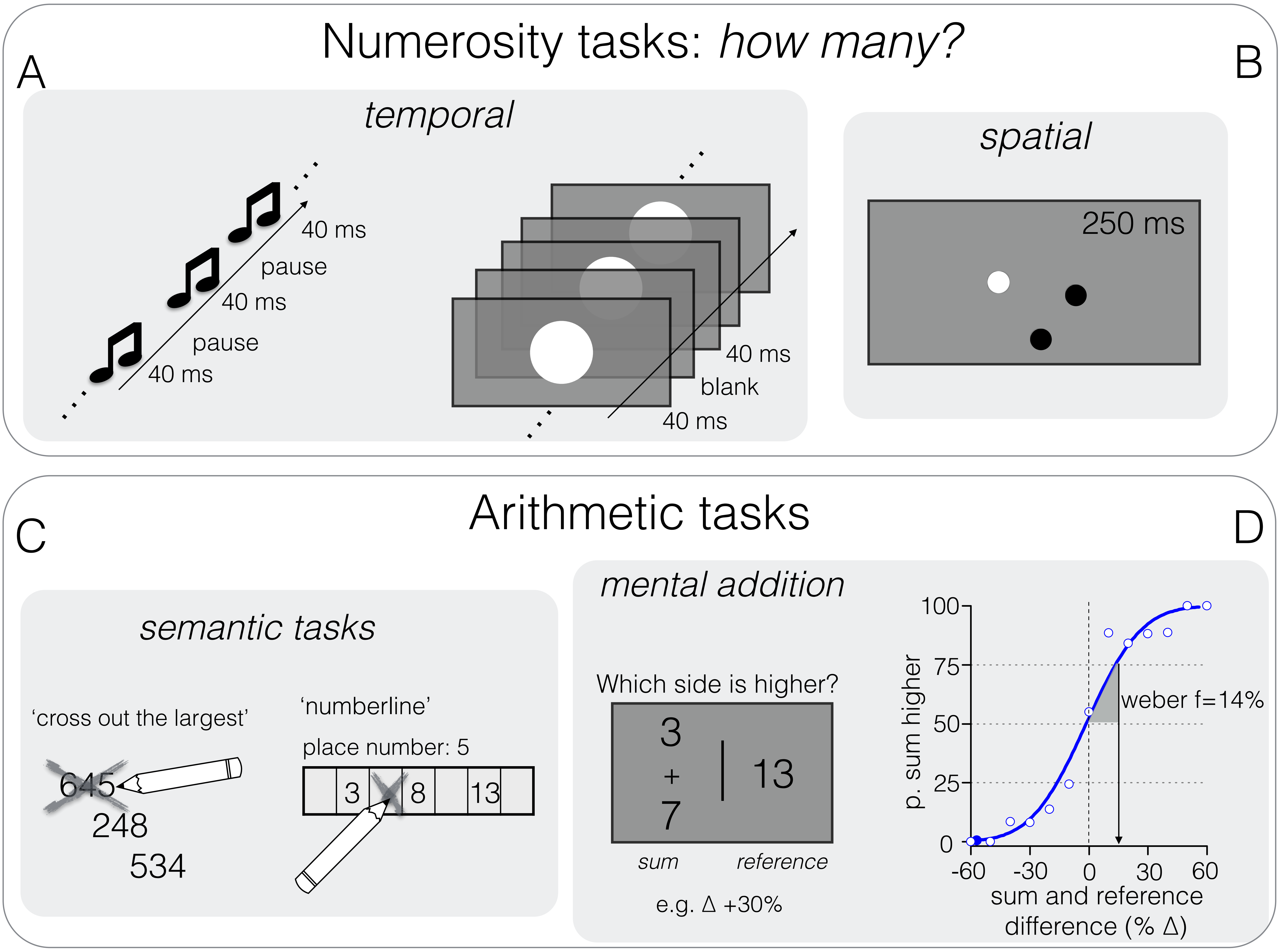
Vetter, P., Butterworth, B., & Bahrami, B. (2008). Modulating attentional load affects numerosity estimation: evidence against a pre-attentive subitizing mechanism. *PLoS One, 3*(9), e3269. doi:10.1371/journal.pone.0003269

Wang, J. J., Halberda, J., & Feigenson, L. (2017). Approximate number sense correlates with math performance in gifted adolescents. *Acta Psychol (Amst), 176*, 78-84. doi:10.1016/j.actpsy.2017.03.014

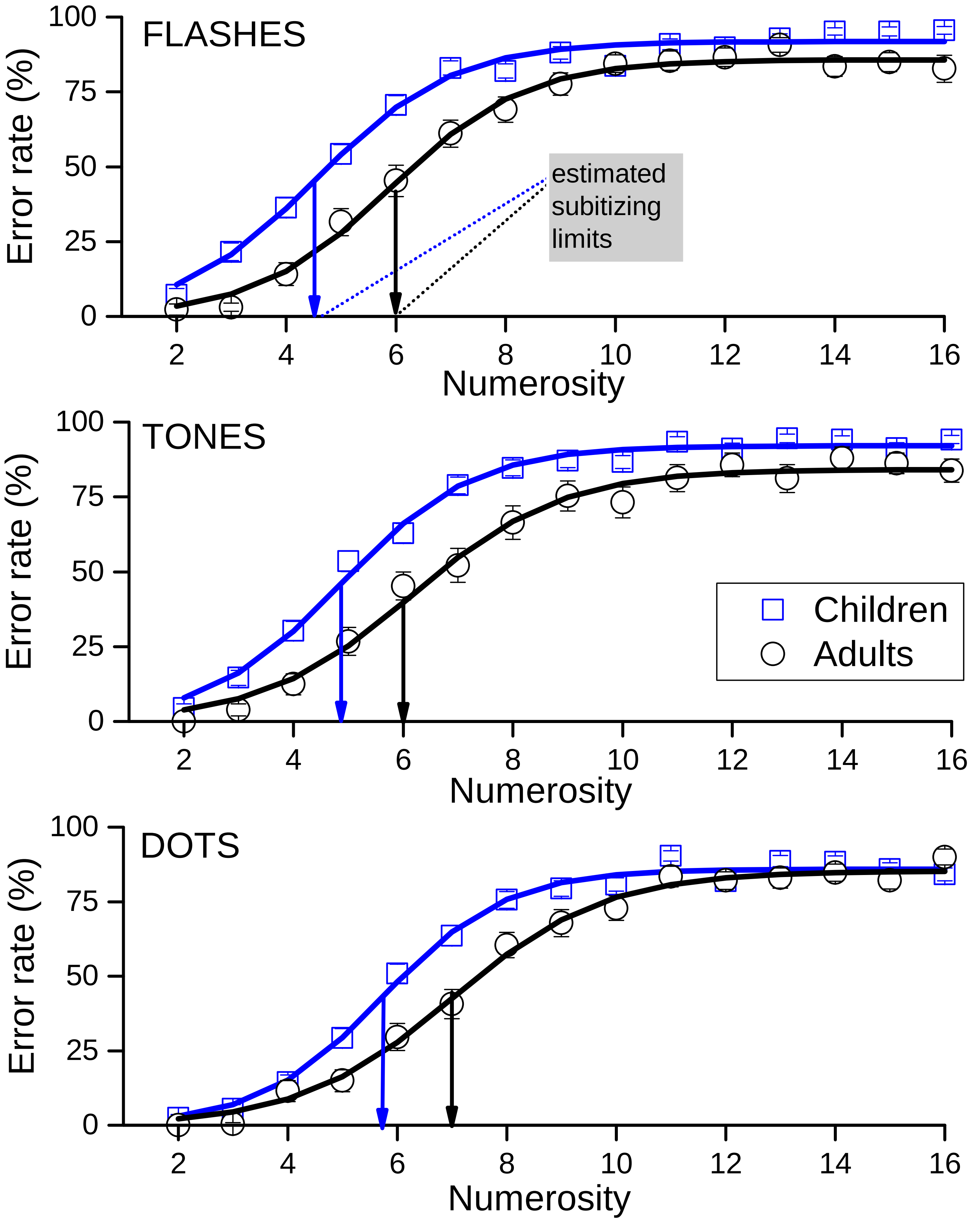
Wetzels, R., & Wagenmakers, E. J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychon Bull Rev, 19*(6), 1057-1064. doi:10.3758/s13423-012-0295-x



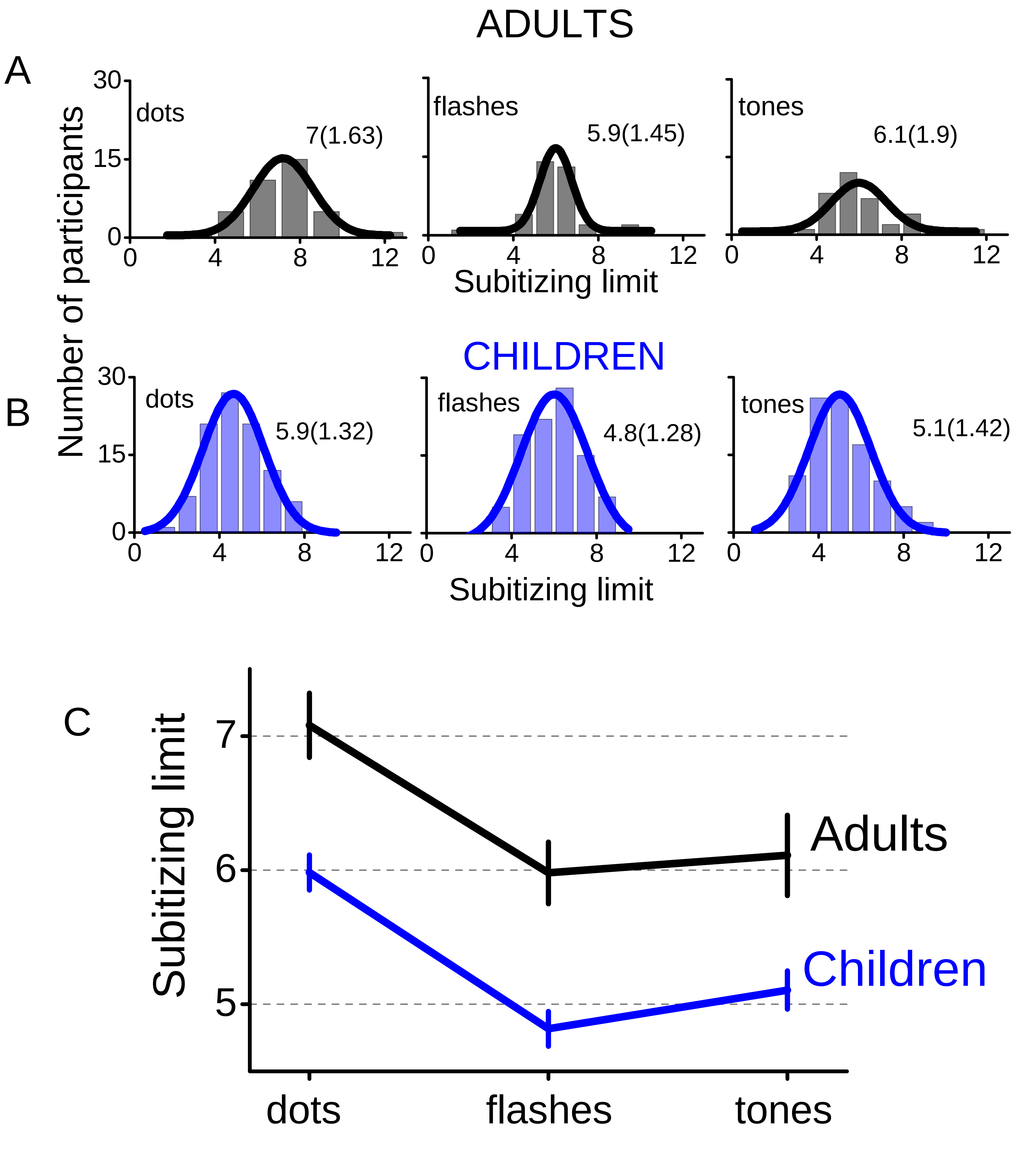
A
sample split on math
frequency
0,2
0,1
0,0
<= 15
percentile
N=14
>= 85
percentile
N=14
-2
-1
0
1
2
Math index
B
Math ability index
distributions in the two tails
frequency
0,30
0,15
0,00
Low
math
High
math
-1,0
-0,5
0,0
0,5
1,0
Math index
Subitizing capacities in the two math distributions
C
0,2
0,1
0,0
Frequency
Tones
Hm=4.65(0.3)
L m=4.64(0.3)
p-value=0.5
2
4
6
8
Dots
D
Hm=5.96(0.4)
L m=5.95(0.4)
p-value=0.49
2
4
6
8
Flashes
E
Hm=4.16(0.3)
L m=4.15(0.3)
p-value=0.50
2
4
6
8
Subitizing capacity



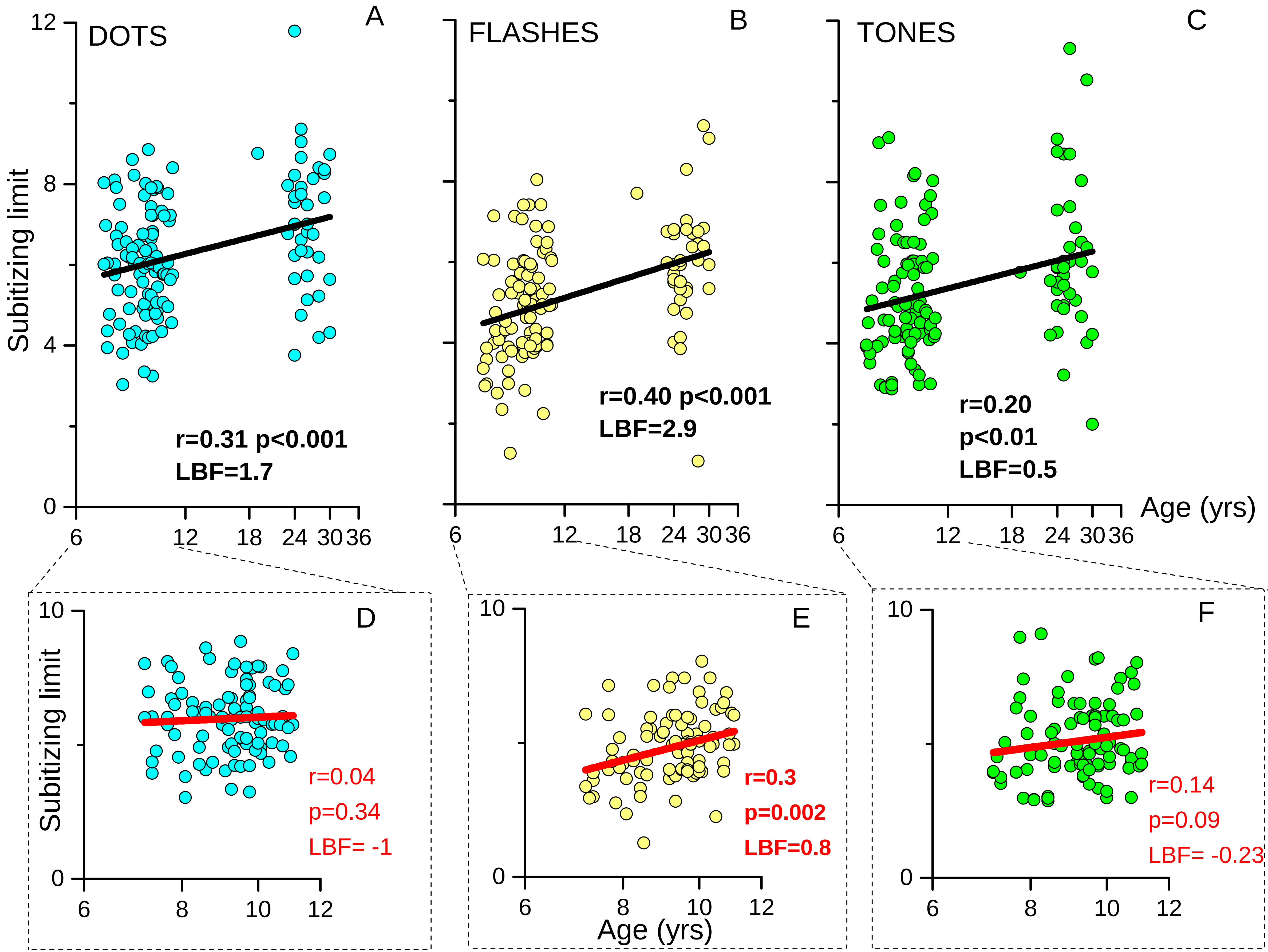
Numerosity tasks: *how many?*
A
B
*temporal*
40 ms
pause
40 ms
pause
40 ms
40 ms
blank
40 ms
*spatial*
250 ms
Arithmetic tasks
C
D
*semantic tasks*
'cross out the largest'
645
248
534
'numberline'
place number: 5
3
8
13
*mental addition*
Which side is higher?
3
+
7
13
*sum*
*reference*
e.g. Δ +30%
p. sum higher
100
75
50
25
0
weber f=14%
-60
-30
0
30
60
sum and reference difference (% Δ)



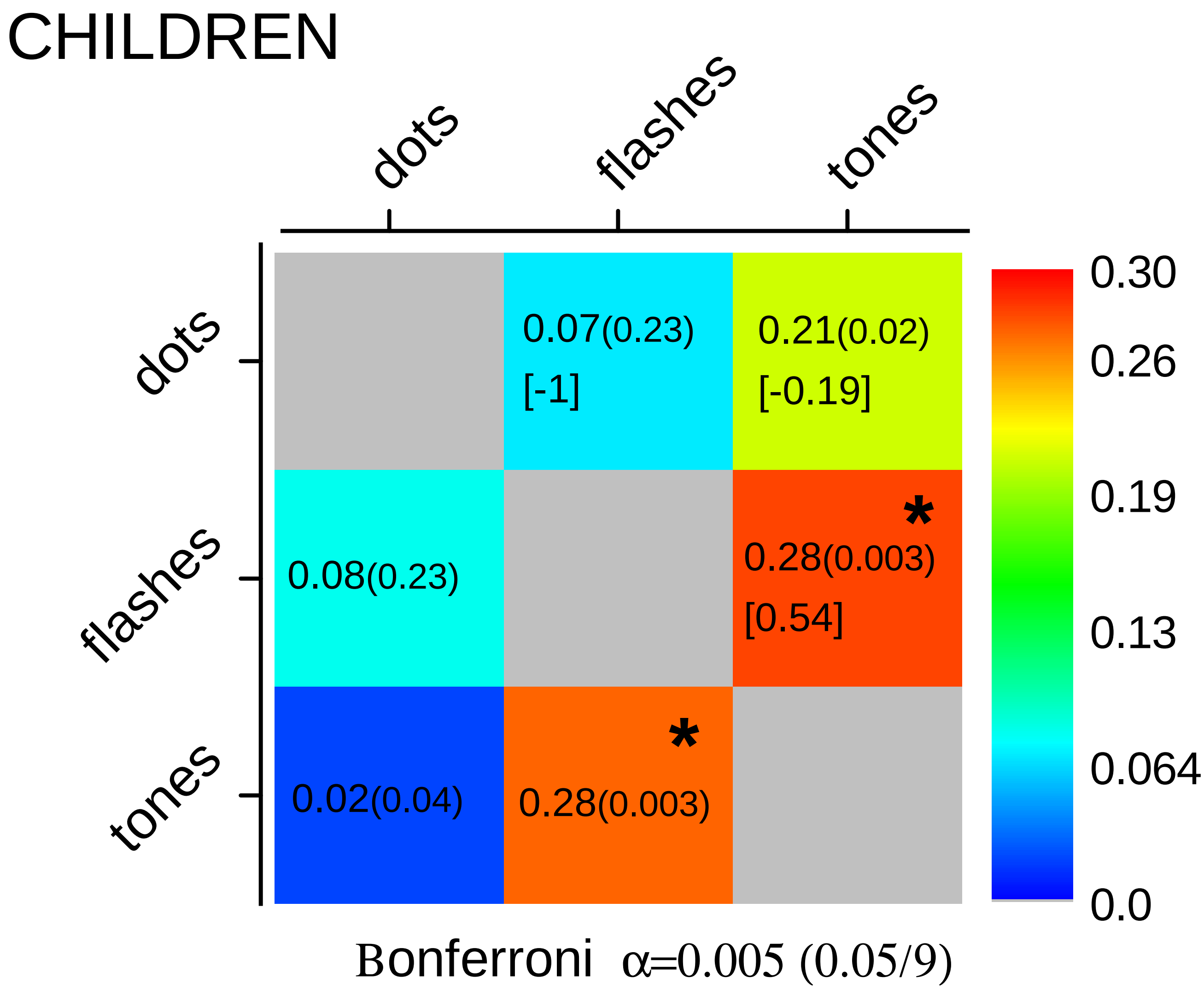
FLASHES
TONES
DOTS
Error rate (%)
100
75
50
25
0
Numerosity
2
4
6
8
10
12
14
16
estimated subitizing limits
Children
Adults



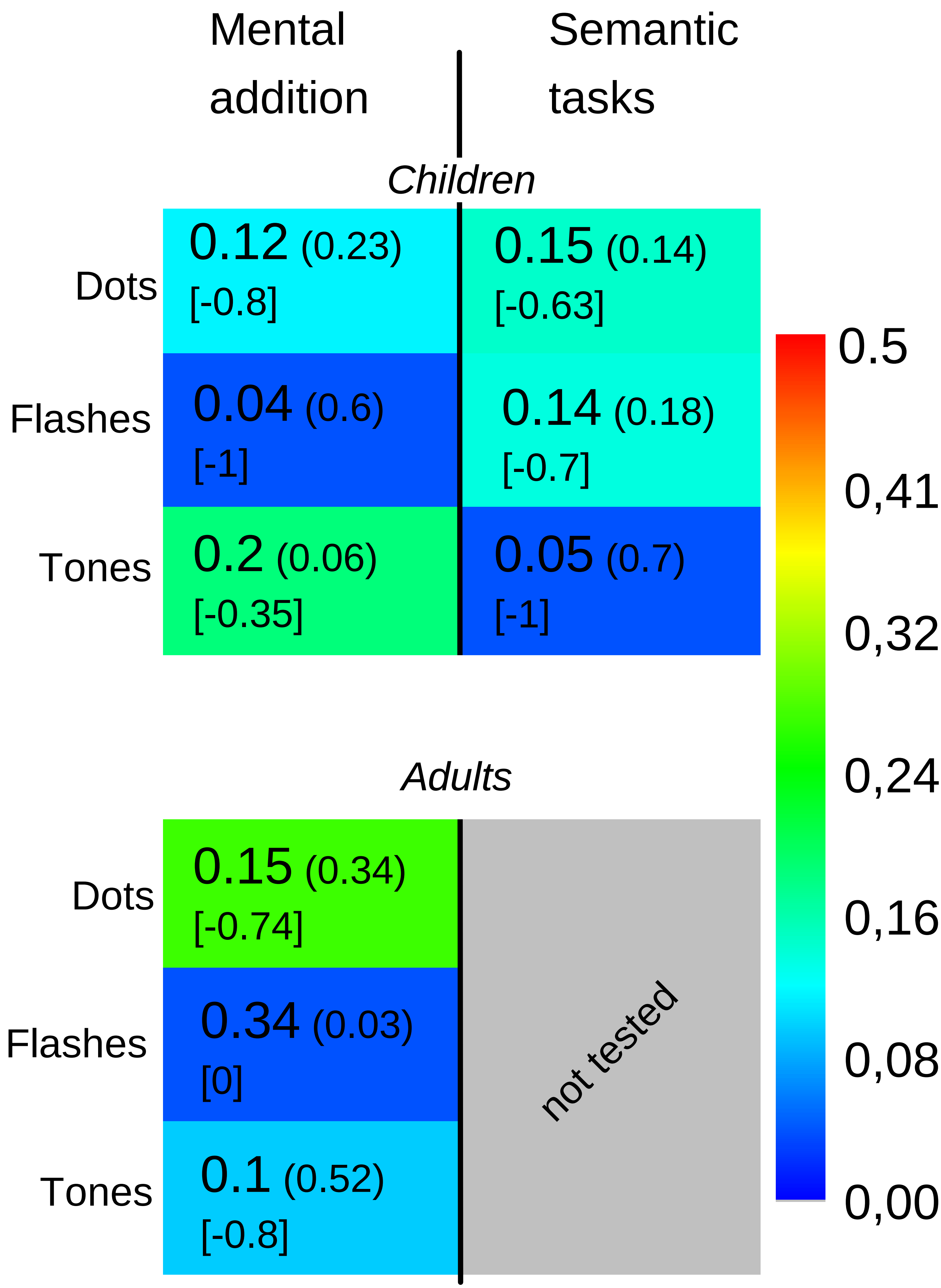
ADULTS
A
Number of participants
dots
7(1.63)
flashes
5.9(1.45)
tones
6.1(1.9)
Subitizing limit
CHILDREN
B
dots
5.9(1.32)
flashes
4.8(1.28)
tones
5.1(1.42)
Subitizing limit
C
Subitizing limit
Adults
Children
dots
flashes
tones



DOTS
A
FLASHES
B
TONES
C
Subitizing limit
12
8
4
0
r=0.31 p<0.001
LBF=1.7
r=0.40 p<0.001
LBF=2.9
r=0.20
p<0.01
LBF=0.5
6
12
18
24
30
36
Age (yrs)
D
E
F
10
0
r=0.04
p=0.34
LBF= -1
r=0.3
p=0.002
LBF=0.8
r=0.14
p=0.09
LBF= -0.23
8
Age (yrs)



CHILDREN
dots
flashes
tones
dots
flashes
tones
0.07(0.23)
[-1]
0.21(0.02)
[-0.19]
0.08(0.23)
*
0.28(0.003)
[0.54]
0.02(0.04)
*
0.28(0.003)
0.30
0.26
0.19
0.13
0.064
0.0
Bonferroni α=0.005 (0.05/9)



Mental addition
Semantic tasks
Children
Dots
0.12 (0.23) [-0.8]
0.15 (0.14) [-0.63]
Flashes
0.04 (0.6) [-1]
0.14 (0.18) [-0.7]
Tones
0.2 (0.06) [-0.35]
0.05 (0.7) [-1]
Adults
Dots
0.15 (0.34) [-0.74]
Flashes
0.34 (0.03) [0]
Tones
0.1 (0.52) [-0.8]
not tested
0.5
0,41
0,32
0,24
0,16
0,08
0,00