



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

## FLORE

# Repository istituzionale dell'Università degli Studi di Firenze

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

Questa è la versione Preprint (Submitted version) della seguente pubblicazione:

*Original Citation:*

Simultaneous and sequential subitizing are separate systems, and neither predicts math abilities / Anobile, Giovanni\*; Arrighi, Roberto; Burr, David C.. - In: JOURNAL OF EXPERIMENTAL CHILD PSYCHOLOGY. - ISSN 0022-0965. - ELETTRONICO. - 178:(2019), pp. 86-103. [10.1016/j.jecp.2018.09.017]

*Availability:*

The webpage <https://hdl.handle.net/2158/1142115> of the repository was last updated on 2018-11-16T15:35:17Z

*Published version:*

DOI: 10.1016/j.jecp.2018.09.017

*Terms of use:*

Open Access

La pubblicazione è resa disponibile sotto le norme e i termini della licenza di deposito, secondo quanto stabilito dalla Policy per l'accesso aperto dell'Università degli Studi di Firenze (<https://www.sba.unifi.it/upload/policy-oa-2016-1.pdf>)

*Publisher copyright claim:*

Conformità alle politiche dell'editore / Compliance to publisher's policies

Questa versione della pubblicazione è conforme a quanto richiesto dalle politiche dell'editore in materia di copyright.

This version of the publication conforms to the publisher's copyright policies.

La data sopra indicata si riferisce all'ultimo aggiornamento della scheda del Repository FloRe - The above-mentioned date refers to the last update of the record in the Institutional Repository FloRe

(Article begins on next page)

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

Giovanni Anobile<sup>1</sup>, Roberto Arrighi<sup>2</sup>, David C. Burr<sup>2,3,4</sup>

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](mailto: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.

# 1 Simultaneous and sequential subitizing are 2 separate systems, and neither predicts math 3 abilities

4  
5 **Keywords:** Numerosity, Subitizing, Numerical cognition, Approximate Number System,  
6 Dyscalculia

## 7 8 9 **ABSTRACT**

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

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

26

27

28

## 29 **INTRODUCTION**

30

31 Although humans are the only species to have evolved a symbolic language-based code  
32 for mathematical concepts, we share with many animals the ability to make rough but rapid  
33 estimates of object numerosity (Agrillo, Miletto Petrazzini, & Bisazza, 2017; Dacke &  
34 Srinivasan, 2008; Dehaene, 2011; Ditz & Nieder, 2015; Nieder, 2016; Petrazzini, Agrillo,  
35 Izard, & Bisazza, 2016; Rugani, Vallortigara, Priftis, & Regolin, 2015). In general,  
36 numerosity estimates are fast and errorless up to 4-6 items; after this range, performance  
37 decreases monotonically, both reaction times (RTs) and accuracy. Kaufman (1949) was  
38 first to coin the term “subitizing”, derived from the Latin *subitus* meaning sudden. Subitizing  
39 can be defined in several ways. Kaufman (1949) measured the subitizing limit as the point  
40 of discontinuity in the distribution of RTs or accuracy. This typically resulted in subitizing  
41 being defined as occurring for stimulus numerosities below 6. Subsequent studies  
42 employed slightly different definitions, all based on performance discontinuities and  
43 resulting in slightly different estimates of the subitizing limit (Arp & Fagard, 2005; Arp,  
44 Taranne, & Fagard, 2006; Ashkenazi, Mark-Zigdon, & Henik, 2013; Burr, Turi, & Anobile,  
45 2010; Camos & Tillmann, 2008; Green & Bavelier, 2003; Olivers & Watson, 2008; Piazza,  
46 Fumarola, Chinello, & Melcher, 2011; Revkin, Piazza, Izard, Cohen, & Dehaene, 2008;  
47 Schleifer & Landerl, 2011). One way is to take the inflexion point of a sigmoid function  
48 fitting estimation’s error or reaction times (Piazza et al., 2011; Revkin et al., 2008). This

49 method tends to overestimate the limit compared with other techniques, but has been  
50 proven to be robust, and well suited to detect inter-individual differences.

51 Subitizing phenomena are thought to be linked to many non-numerical capacities like  
52 attention, working memory and object tracking. All these mechanisms are capacity-limited  
53 and interact with each other, making it difficult to assess their individual contribution. For  
54 example, simultaneous subitizing is heavily affected by the deployment of simultaneous  
55 and temporal attentional resources (Burr et al., 2010; Olivers & Watson, 2008; Vetter,  
56 Butterworth, & Bahrami, 2008) as well as visual working memory (Melcher & Piazza, 2011;  
57 Piazza et al., 2011). All these results clearly suggest the existence of partially shared  
58 mechanisms. The so-called "Object Tracking System" (OTS), the process involved in  
59 identifying, representing and tracking objects through time and space is, like subitizing,  
60 strongly dependent on attentional resources (Arrighi, Lunardi, & Burr, 2011; Pylyshyn &  
61 Storm, 1988). However, there are clear differences between these processes. For  
62 example, OTS capacity has been found to be adult-like at one year of age (Piazza, 2010),  
63 while spatial working memory continues to develop until 6/7 years (Cowan, Morey,  
64 AuBuchon, Zwillig, & Gilchrist, 2010). Also OTS capacity measured by a visual multiple  
65 object tracking task is affected by visual but not auditory attentional deprivation (Arrighi et  
66 al., 2011), while visual subitizing strongly suffers from cross-modal (visual, auditory and  
67 haptic) dual tasks (Anobile, Turi, Cicchini, & Burr, 2012). Some evidence also suggests a  
68 link between individual differences in working memory (Bull & Scerif, 2001; De Smedt et  
69 al., 2009; Passolunghi & Siegel, 2004; Raghubara, Barnesb, & Hechtb, 2010; Toll,  
70 Kroesbergen, & Van Luit, 2016), attention (Ashkenazi & Henik, 2012; Askenazi & Henik,  
71 2010; Steele, Karmiloff-Smith, Cornish, & Scerif, 2012) as well as OTS (Anobile, Stievano,  
72 & Burr, 2013) and math performance. In brief, while subitizing has been extensively  
73 studied, the underlying mechanisms, and how it relates to other cognitive capacities are  
74 still unclear.

75 Moreover, while subitizing has been extensively studied, most studies concentrate on  
76 spatial arrays of simultaneous visual stimuli; and it is still not clear whether subitizing  
77 generalizes to other sensory modalities. Very few studies have investigated subitizing for  
78 temporal sequences, or in sensory modalities other than vision. The available data  
79 suggest that adults may subitize auditory sequences (Camos & Tillmann, 2008; Repp,  
80 2007), but not simultaneously played sounds (McLachlan, Marco, & Wilson, 2012). In  
81 vision, only one study has described subitizing for temporal sequences (Camos &  
82 Tillmann, 2008), while more evidence has been provided for the existence of subitizing  
83 sequences of haptic stimuli (Ferrand, Riggs, & Castronovo, 2010; Gallace, Tan, Haggard,  
84 & Spence, 2008; Plaisier, Bergmann Tiest, & Kappers, 2009; Plaisier & Smeets, 2011;  
85 Plaisier, Tiest, & Kappers, 2010; Riggs et al., 2006). All these studies involved adult  
86 participants, and none took into account between-task correlations.

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

100 8.5 years old showed subitizing difficulties (Desoete & Grégoire, 2006). In summary, the  
101 evidence for a link between subitizing and math is somewhat variable.

102 The numerical range above subitizing is termed the “estimation” range, thought to reflect  
103 the action of the “Approximate Number System (ANS)” (Feigenson, Dehaene, & Spelke,  
104 2004). The ANS is a generalized system, encoding and integrating information across  
105 sensorimotor domains, including vision, audition and action, and stimulus formats, both  
106 simultaneous spatial ensembles and temporal sequences (Anobile, Arrighi, Togoli, & Burr,  
107 2016; Arrighi, Togoli, & Burr, 2014; Izard, Sann, Spelke, & Streri, 2009). ANS precision  
108 correlates with current and future children math abilities along the entire spectrum of  
109 mathematical abilities, from low-math to math-gifted children (Piazza et al., 2010; Wang,  
110 Halberda, & Feigenson, 2017), including ‘average’ math skilled children.

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

116

## 117 **METHODS**

118

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

125

126 **General procedures.** Stimuli were generated and presented with Matlab 8.1 using  
127 PsychToolbox routines (Brainard, 1997) on a 17-inch LG touch screen monitor with 1280 X  
128 1024 resolution at refresh rate of 60 Hz. Each participant was tested in two separate  
129 sessions (usually within the same week), lasting around 1 hour each. Math abilities were  
130 measured by a paper and pencil test (only children), and by a computerized digit  
131 summation task (children and adults). All participants also performed a non-verbal  
132 reasoning task (Raven matrices). Math skills and nonverbal reasoning were measured at  
133 the end of the first session, perceptual tasks were administered in a pseudorandom order  
134 between participants. This study is based on a new analysis of a set of data collected for  
135 other purposes (Anobile et al., 2017). The experimental methods used here were the  
136 same as in the previous study but here we focused on the subitizing range by not  
137 excluding numerosities  $\leq 4$  from the analyses.

138

139 **Numerosity estimation.** Visual stimuli were either ensembles of 0.5° diameter dots half-  
140 white, half-black in order to balance luminance across numerosities (in case of odd  
141 numbers the one excess dot was randomly assigned to white or black), presented  
142 simultaneously for 250 ms within a virtual 16° diameter region, or sequences of flashes  
143 (sharp-edged white discs of 90 cd m<sup>-2</sup> and 5° diameter) presented in a pseudo-random  
144 order within a 2 second interval (Figure 1 A and B). In the sequential presentation, each  
145 flash lasted 40 ms with the constraint that two pulses could not fall within 40 ms of each  
146 other. All visual stimuli were presented centrally, with subject viewing distance set at 57  
147 cm, on a grey background of 40 cd/m<sup>2</sup>. Precision for estimates of sequential numerosity  
148 was also investigated in audition, with 500 Hz pure tones ramped on and off with 5 ms  
149 raised-cosine ramps, presented with an intensity of 80 dB (at the sound source) and  
150 digitized at a rate of 65 kHz. Sounds were presented through high-quality headphones  
151 Microsoft lifechat LX-3000, and perceptually localized in the middle of the head. In all

152 conditions, the numerosity range was 2-18 and subjects were asked to verbally report the  
153 number of perceived stimuli, which was recorded by the experimenter via a computer  
154 keyboard. The testing phase was preceded by a training session of 17 trials (not included  
155 in the main analyses). During training, all numerosities were randomly presented, and  
156 feedback was provided displaying the actual numerosity on the monitor screen. The aim of  
157 feedback was to calibrate participants' judgements (mainly young children) to have all  
158 estimates within the numerical range without aberrant responses (for a similar procedure  
159 see Revkin et al., 2008). After training had been completed, the testing phase started with  
160 a block of 51 trials (3 repetitions for each numerosity), with no feedback. In total each  
161 participant performed 204 trials. Test numerosity ranged from 2 to 18, but we computed  
162 error rates and fitted (see later) only the range 2-16 to avoid edge effects. Average  
163 temporal rates for both flashes and sounds stimuli were 640 ms (min 110, max 1180), 500  
164 ms (min 130, max 900) and 400 ms (min 140, max 650) respectively for numerosity two,  
165 three and four. As the counting speed for numbers in Italian primary school children is  
166 around 600/800 milliseconds per number, and the stimulus sequences were not regular  
167 but jittered in time, it is unlikely that children were able to serially count the stimuli.

168

169 **Semantic skills.** Two paper and pencil tasks were administered (Figure 1 C): 1) mark the  
170 largest numbers in a set of three (one to five digits, 36 trials); and 2) mark where a number  
171 should be placed (four possible positions among three other numbers, one to six digits, 18  
172 trials). These tasks were extracted from an Italian standardized battery suitable for children  
173 from 8 to 13 years old, not suitable for adults (Biancardi, Bachmann, & Nicoletti, 2016).  
174 They are thought to tap the semantic component of numeracy (Dehaene, Piazza, Pinel, &  
175 Cohen, 2003), and have been demonstrated to be good predictor of children numerosity  
176 discrimination thresholds (Anobile et al., 2013; Cicchini, Anobile, & Burr, 2016; Piazza,  
177 2010). Again, accuracy and speed were measured (as the sum of errors and time in

178 minutes required to complete the three tasks). Similarly to the mental calculation task (see  
179 below) we measured two separate z-scores for speed and accuracy and computed a  
180 performance-combined index averaging the two z-scores (same technique exploited by  
181 Anobile et al., 2017; Cicchini et al., 2016).

182

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

198 Not every trial where RTs exceeded the threshold were eliminated from the analysis, as  
199 we applied a within subject cut-off: for each participant we measured the average reaction  
200 times (across trials) and eliminated those higher or lower than 3 standard deviations. The  
201 total number of eliminated trials was 38 (1.1%) for adults and 80 (1.4%) for children. The  
202 proportion of 'sum higher' was plotted against the percentage difference between the sum  
203 and the single digit. We fitted the data with a cumulative Gaussian error functions. The

204 percentage difference needed to move from 50% to 75% correct responses provided an  
205 “mental additional discrimination threshold”. This is logically equivalent to the Weber  
206 fraction usually measured for numerosity discrimination tasks, and could be interpreted as  
207 the amount of noise present in the mental addition process (see Figure 1 D). Similarly to  
208 previous studies (Anobile et al., 2017; Cicchini et al., 2016), we computed for each  
209 participant two separate z-scores: one for precision (Weber fraction) and the other for  
210 response speed (RT). Z-scores were measured using the mean and standard deviation of  
211 the participant grade class (from second to fifth grade). For adults we used the mean and  
212 standard deviation of the entire group. Finally, for each participant we computed a  
213 performance-combined index averaging the two z-scores. A previous study demonstrated  
214 that children’s performance on this task is a good predictor of their numerosity estimation  
215 precision of simultaneous dot arrays (Weber Fraction), for numerosity above the subitizing  
216 range (Anobile et al., 2017).

217

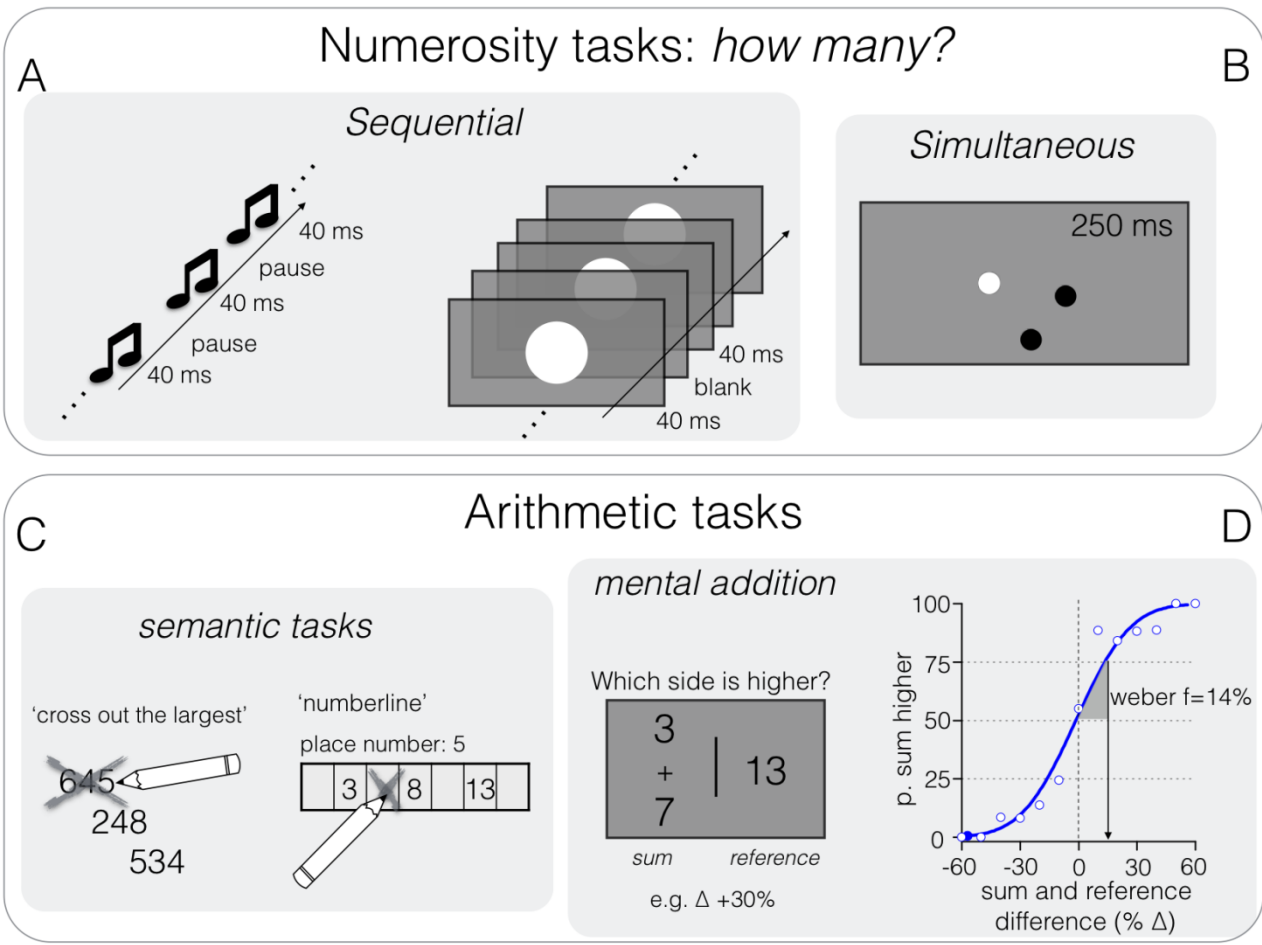
218 **Data analysis.** Following previous works (Piazza et al., 2011; Revkin et al., 2008), we  
219 fitted error rates with sigmoid functions and defined the subitizing range as the inflection  
220 point of the function. We performed this procedure separately for each participant as well  
221 as for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

225 Correlation analyses were performed by both zero-order and partial Pearson correlations  
226 procedures. Statistical significance was indexed by p-values and Bayes factor (Wetzels &  
227 Wagenmakers, 2012). Bayes factor is the ratio of the likelihood probabilities of the two  
228 models, that a correlation  $s$  quantifies the ratio of the likelihood probabilities  $H_1/H_0$ , where  
229  $H_1$  is the likelihood of a correlation between the two variables, and  $H_0$  the likelihood that

230 the correlation does not exist. By convention, a Log Bayes Factor (LBF) greater than 0.5 it  
231 is considered substantial evidence in favour of the existence of the correlation, and  $LBF <$   
232  $-0.5$  substantial evidence in favour of it not existing. Absolute values of LBF greater than 1  
233 are considered strong evidence, and values greater than 2 are considered decisive.  
234 Missing values were left empty and data excluded with pairwise deletion method.

235

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



256

257

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

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

270

271

272

## 273 **RESULTS**

274

275 Task reliability. Table 1 reports split-half reliability levels (Pearson's  $r$ ) for all the tasks for  
276 subject-by-subject analyses. Indexes were all reasonably good, ranging from 0.57 to 0.97.

277 Regarding split-half subitizing reliability measures on pooled data, we found no statistically  
278 significant differences between average capacities calculated from the two data halves in  
279 both groups of children and adults (Table 2).

280

281

Table 1 – Split-half reliability indexes for children (C) and adults (A). Errors reflect standard errors.	
Tasks	Pearson's $r$
Dots	C: $0.64 \pm 0.18$ A: $0.65 \pm 0.19$
Flashes	C: $0.57 \pm 0.24$ A: $0.64 \pm 0.20$
Tones	C: $0.68 \pm 0.13$ A: $0.74 \pm 0.17$
Mental	Precision (Weber Fraction)

addition	C: 0.58±0.18
	A: 0.75±0.08
	Speed (Reaction Times)
	C: 0.97±0.006
	A: 0.95±0.01

282

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.

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

283

284 **Goodness of fit**

285 The sigmoid fits describe well the data (Figure 1), with good coefficients of determination  
286 ( $R^2$ ). Indeed, between participants average  $R^2$  for children were 0.74 (SD 0.14, min 0.32),  
287 0.73 (SD 0.17, min 0.33), and 0.76 (SD 0.37, min 0.37), for dots, flashes and tones  
288 estimations, respectively.  $R^2$  fits for adults were on average 0.79 (SD 0.14, min 0.25), 0.76  
289 (SD 0.17, min 0.42), and 0.75 (SD 0.18, min 0.38) for dots, flashes and tones estimations  
290 respectively. Some participants had at least one condition in which the  $R^2$  was too poor to  
291 reliably estimate the subitizing limit. Similar Piazza et al's (2010), we adopted criterion of  
292 eliminating subjects with  $R^2 < 0.25$ . Ten children had a total of twelve low  $R^2$  fit values (two  
293 fit for dots stimuli, six with flashes, four with sounds). Also two adults had poor  $R^2$  (one for  
294 flash stimuli and one for sounds).

295

296

### 297 **Subitizing limits in children and adults**

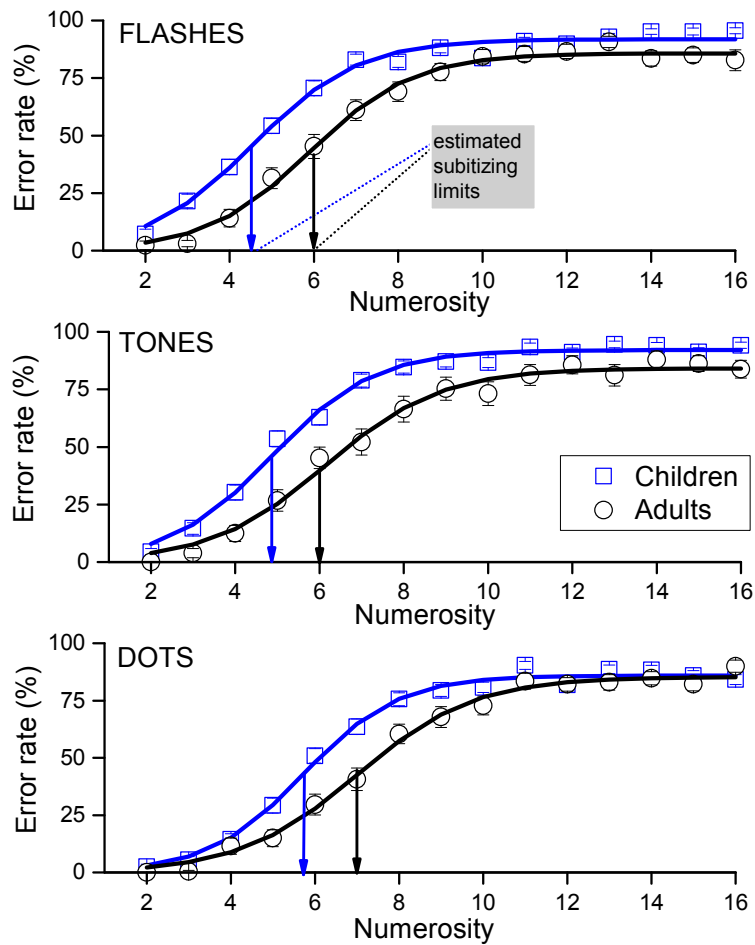
298 We measured error rates (symbols on Figure 2) for estimating numerosity of dots arrays,  
299 sequences of flashes and auditory (tones) events. We fitted errors with sigmoid functions  
300 and took the inflection point as an index of subitizing limit. Figure 2 shows averaged  
301 results: all conditions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

304

305

306

307



308

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

314

315 The same fitting procedure was applied separately to each participant. Figures 3 A and B  
 316 show the frequency distributions of subitizing limits across participants. On average, limits  
 317 peaked around five/six items (overall averages, pooling together data for all stimuli, were  
 318 5.26 and 6.3 for children and adults respectively), a range often reported in the literature  
 319 (for similar value in case of dots see Kaufman & Lord, 1949). More importantly, the  
 320 distributions show much inter-subject variability. This replicates previous findings (Piazza

321 et al., 2011), and suggests that the variance is large enough to run correlational analyses  
 322 (described in the next paragraph).

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

335  
 336  
 337  
 338  
 339  
 340

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

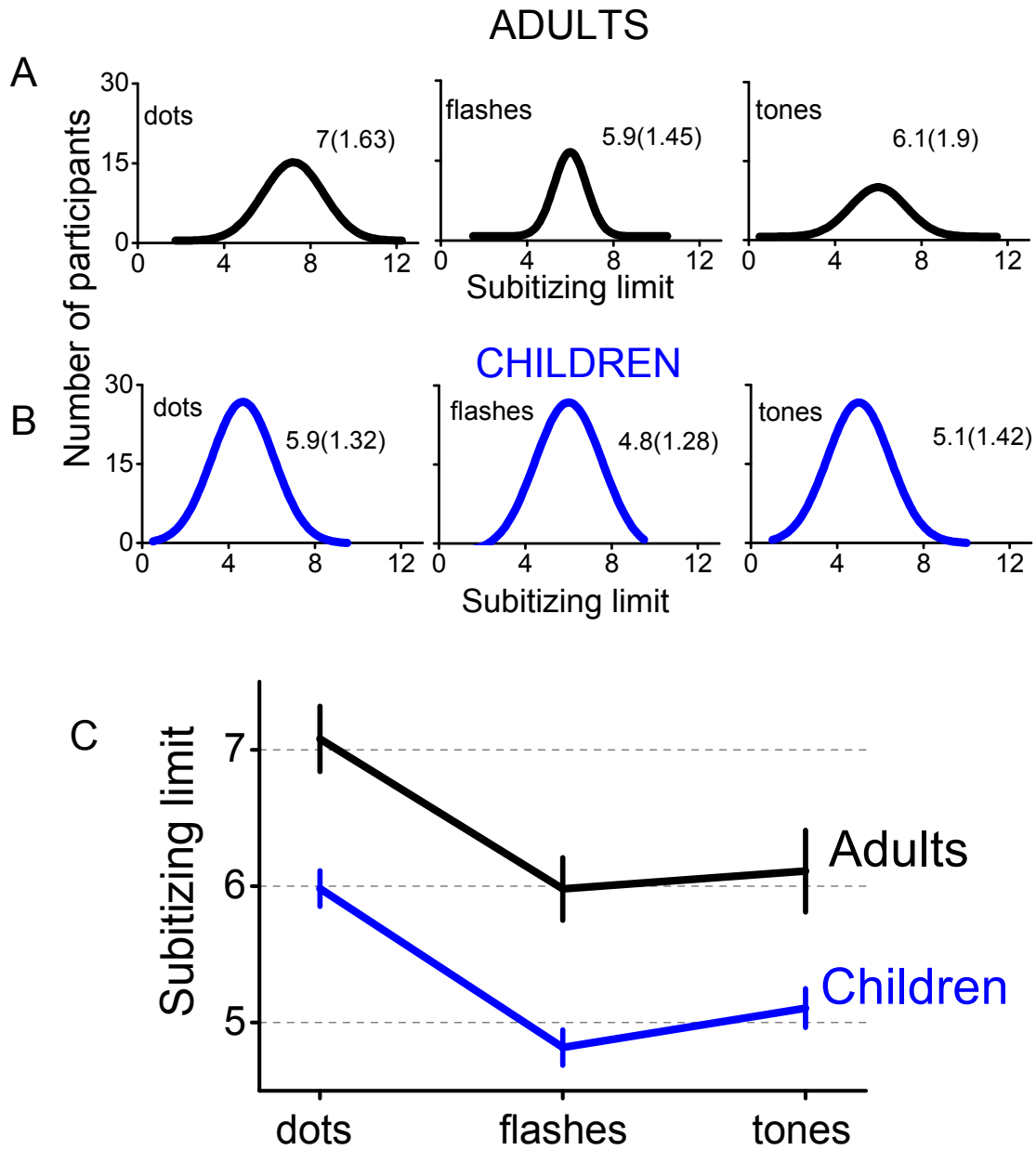
Stimuli	Difference	95% CI of the difference	
		Low	High
Dots vs Flashes	C: 1.16***	C: 0.69	C: 1.63
	A: 1.09**	A: 0.16	A: 2.03
Dots vs Tones	C: 0.87***	C: 0.41	C: 1.34
	A: 0.97*	A: 0.04	A: 1.9
Tones vs Flashes	C: 0.28 n.s.	C: -0.75	C: 0.17
	A: 0.13 n.s.	A: -0.8	A: 1.06

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

341

342

343



344

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

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

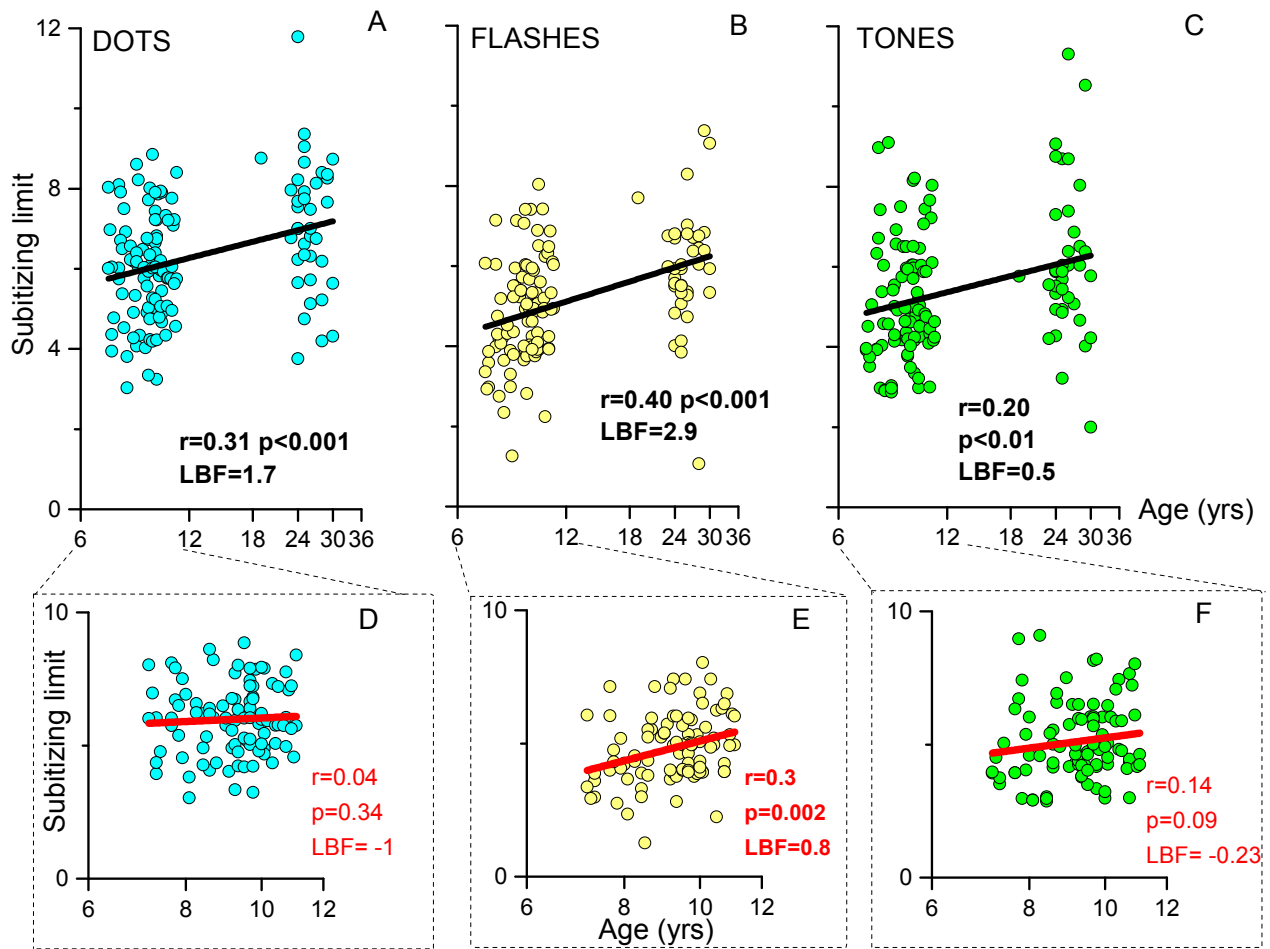
351

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

365

366

367



368

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

371

372

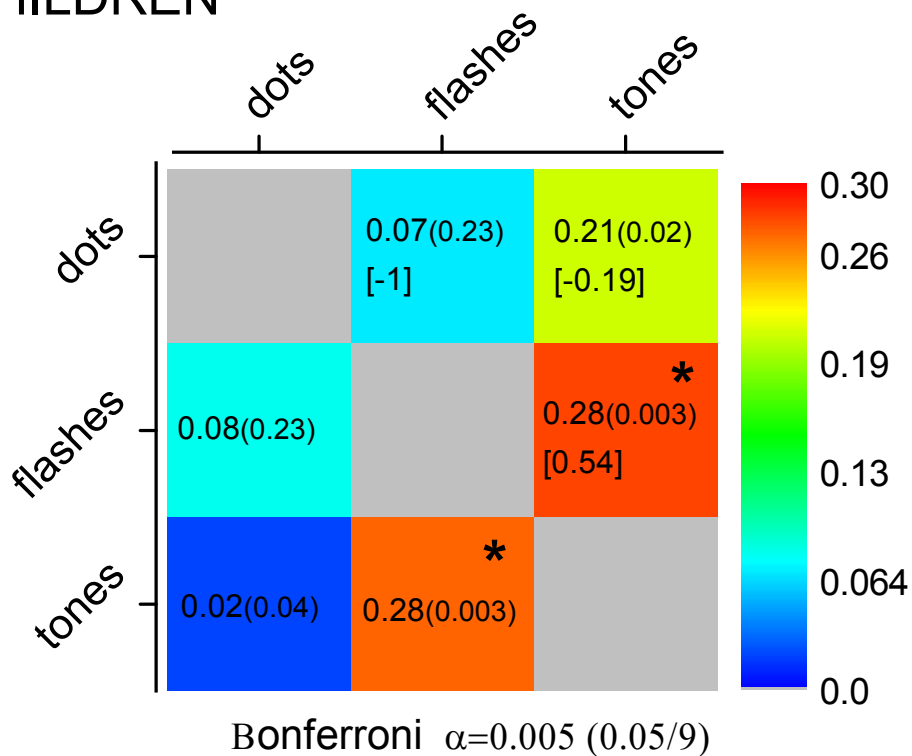
373 **Correlation between simultaneous and sequential subitizing**

374 The results so far show that both adults and children can subitize simultaneous and  
 375 sequential stimuli. We also found that simultaneous subitizing outperforms sequential  
 376 subitizing, regardless of the sensory modality of the stimuli, and that subitizing capacities  
 377 for different kind of stimuli develop differently during childhood, suggesting different  
 378 systems. Here we investigated this possibility further by correlating simultaneous and  
 379 sequential subitizing limits. The results show that child subitizing limits for sequential  
 380 stimuli positively correlate between each other: children with higher subitizing limits for

381 sequential sequences of visual events also have higher limits for sequences of tones  
 382 ( $r=0.28$ ,  $p= 0.003$ ,  $LBF=0.54$ , Table 4). However, subitizing for simultaneous stimuli did not  
 383 correlate with any of the sequential conditions (dots vs flashes  $r=0.07$ ,  $p=0.23$ ,  $LBF = -1$ ;  
 384 dots vs tones  $r=0.21$ ,  $p=0.02$ ,  $LBF = -0.19$ ). The positive significant correlation between  
 385 sequential stimuli was robust, as it remains significant even when the effect of age and  
 386 non-verbal IQ were simultaneously controlled for ( $r_p=0.28$ ,  $p= 0.006$ , Table 4, below  
 387 diagonal). No significant correlation was found for adult participants (dots vs flashes  $r=$   
 388  $0.28$   $p=0.08$ ,  $LBF = -0.28$ ; dots vs tones  $r=0.27$ ,  $p=0.09$ ,  $LBF = -0.92$ ; flashes vs tones  
 389  $r=0.07$ ,  $p=0.64$ ,  $LBF= -0.89$ ). These analyses are in line with those previously reported in  
 390 this manuscript to support the idea of two different systems for simultaneous and  
 391 sequential subitizing.

392

## CHILDREN



393

394

395

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

400

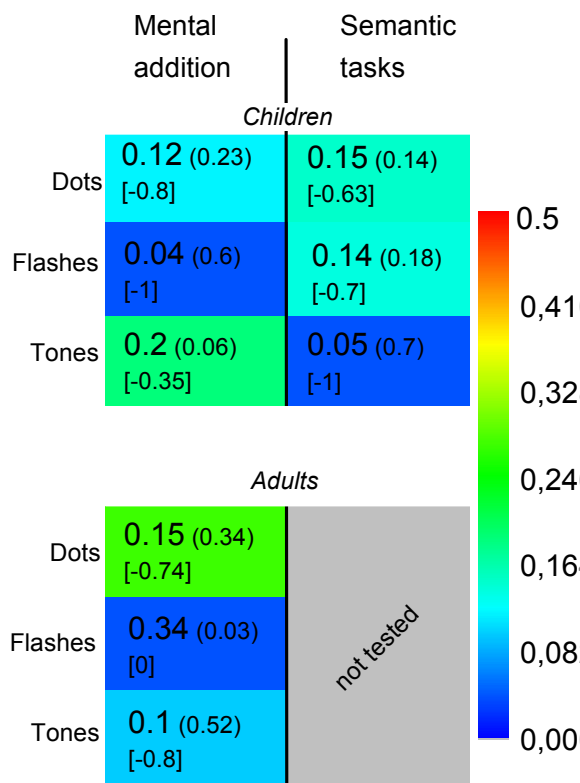
#### 401 **Correlation between subitizing limits and mathematical abilities**

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

414

415

416



417

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

422

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

430

**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  $\alpha=0.008$  (0.05/6).

432

433

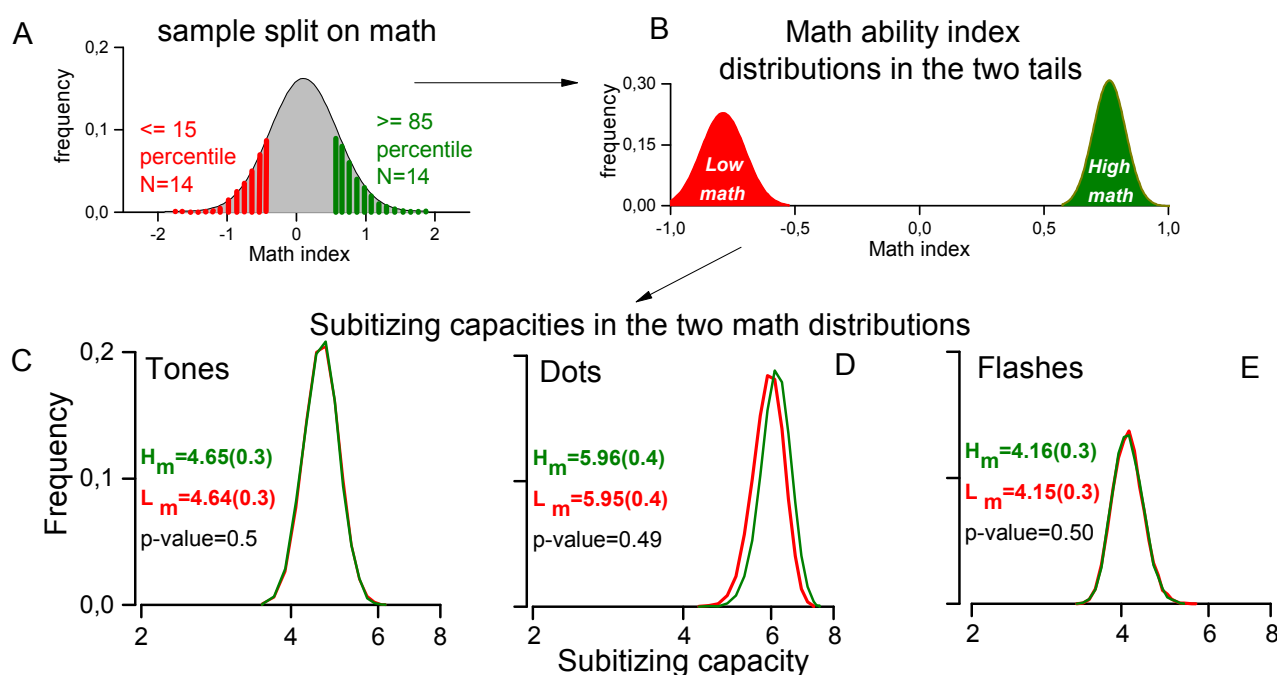
434 Correlational analyses may have missed potential clumping in the data and thus obscured  
 435 the presence of potential distinct subpopulations in the sample. To explore this possibility,  
 436 we consider child data (the larger sample size) and ran a two-step cluster analysis by  
 437 considering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

447

448 We then applied a more extreme approach: we split the child sample into two sub-  
449 samples, collapsing higher and lower math skilled subjects. Because the performance on  
450 two symbolic math tasks correlated well between each other ( $r=0.43$ ,  $p<0.001$ ), we built a  
451 summary math index by averaging the z-scores of the two tests, and used this index value  
452 to compute child math percentiles. Those above the 85<sup>th</sup> percentile were assigned to the  
453 “high-math” group (N14, average z-score 0.76), those below the 15<sup>th</sup> assigned to the “low-  
454 math” group (N14, average z-score  $-0.79$ ). With a non-parametric sample-with-  
455 replacement bootstrap technique (10000 iterations) we built the two-math distributions  
456 shown in Figure 5: on each iteration math z-scores were resampled (with replacement)  
457 separately for each math group, and the average z-scores computed. At the same time we  
458 calculated subitizing limits, separately for each stimulus condition and math group. Figure  
459 5 (C, D and E) shows the average subitizing frequency distributions for the two math  
460 groups (red “low”, green “high”). Those distributions largely overlap with virtually no  
461 subitizing advantage for high-math children for any of the stimulus conditions. We  
462 statistically computed the difference between those distributions, counting the times that,  
463 on each of the 10000 iterations, the difference between the averaged subitizing capacities  
464 were higher than zero (one-tail p-value). All p-values (reported in Figure 5 C, D, E) were  
465 near 0.5, showing clearly that they were not statistically different, robustly reinforcing our  
466 finding of no correlation between subitizing and math capacities.

467 We repeated the analysis with even more conservative criteria, considering only the 5%  
468 tails of the math distribution (5-95 percentile, samples size was  $N=5$  and  $N=6$  for  $\leq 5\%$  and  
469  $\geq 95\%$  respectively) or the 10% (10-90 percentile,  $N=11$  each group). Even in these cases  
470 none of the subitizing measures differed between groups (all  $p\approx 0.5$ ). We then checked  
471 whether an automatized analysis might have identified the same group of subjects than  
472 our custom procedure by running a cluster analysis based on math abilities (overall index).  
473 The obtained dendrogram revealed a maximum distance between a small group of

474 participants (with particularly low scores on math ability) and the rest of the participants.  
 475 Interestingly, this group consisted of the very same participants included in the 5<sup>th</sup>  
 476 percentile of math proficiency distribution (that as shown above did show a subitizing  
 477 range rather identical to all other participants, even to those in the 5<sup>th</sup> percentile of the  
 478 highest math scores.  
 479



480

481

482 **Figure 5. Subitizing limits for “low” and “high” math children.** A) Frequency  
 483 distribution of child math scores, with tails (14 each) falling in the selected percentile  
 484 ranges: 1) ≥ 85, green, “high math” and 2) ≤ 15, red, “low math”. B) Child math scores  
 485 frequency distributions inside the two percentile samples. C, D and E) Subitizing limit  
 486 distributions separated for the two math samples (H<sub>m</sub> = high math, green; L<sub>m</sub> = low math,  
 487 red). Values reflect: average and SD (in brackets) subitizing capacities and associated  
 488 one tail p-values.

489

490

## 491 **DISCUSSION**

492

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

504

### 505 **Subitizing and math abilities**

506 **The null correlation between subitizing and math scores** may sound counterintuitive,  
507 especially for simultaneous subitizing (**dots**), but some issues are worth considering. First,  
508 the literature linking symbolic math and subitizing limits is not very solid. The most robust  
509 studies describing a link between math abilities and subitizing encompass patients with  
510 math impairments associated with neurological disorders such as cerebral palsy (Arp &  
511 Fagard, 2005; Arp et al., 2006), Turner's syndrome (Bruandet, Molko, Cohen, & Dehaene,  
512 2004), Williams and Down syndrome (Paterson, Girelli, Butterworth, & Karmiloff-Smith,  
513 2006), Gerstmann's syndrome (Cipolotti, Butterworth, & Denes, 1991; Lemer, Dehaene,  
514 Spelke, & Cohen, 2003). Although they are informative, these studies suffer from the  
515 caveat that non-numerical deficits associated with those neurological disorders may have  
516 impacted subitizing.

517 Another important point to consider is that we discouraged serial counting by using a very  
518 fast presentation time (for dots) or high temporal rate (for flashes and sounds), and did not  
519 require speeded responses. However, our error rate was very similar to that commonly  
520 documented in the literature, and measures of reliability were always high. In the light of  
521 these data we can reasonably assume that the methods we applied to measure subitizing  
522 are sensitive and robust. We quantified the likelihood of this null correlation by means of  
523 Bayes factor. Log Bayes factors for correlations between subitizing limits and math skill  
524 were all clearly negative and mainly near  $-1$  (Table 5, square brackets), indicating strong  
525 evidence in favour of the null hypothesis of zero correlation (Wetzels & Wagenmakers,  
526 2012). It is also worth noting that object serial counting speed has been found to be a  
527 good and stable marker of dyscalculia (Gray & Reeve, 2014; Reeve, Reynolds,  
528 Humberstone, & Butterworth, 2012), leaving open the interesting possibility that [leaving](#)  
529 [open the interesting possibility that the link between math and subitizing occurs only when](#)  
530 [counting is used.](#)

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

542 controls. Here, as in other studies (Green & Bavelier, 2003; Piazza et al., 2011), we  
543 focused on this latter parameter (point of discontinuity).

544 In this study we focused on arithmetic tasks that particularly tap into rote memorization and  
545 semantic skills, abilities that have been previously found to correlate with precision in  
546 estimating numerosity above the subitizing range. However, we are aware that these  
547 results do not exclude the possibility that other arithmetic tasks might instead correlate  
548 with subjects' ability to subitize. As mentioned above, no causal studies investigated the  
549 link between math and subitizing. Curiously the only cognitive capacity that was causally  
550 related to subitizing was not numerical. In their seminal paper Green and Bavelier (2003)  
551 showed that boosting visual attention capacities by playing action video games enlarged  
552 the subitizing capacity of adults. In line with that, depriving attention greatly degrades  
553 subitizing, far more than estimation (Anobile et al., 2012; Burr, Anobile, & Turi, 2011; Burr  
554 et al., 2010; Pagano, Lombardi, & Mazza, 2014; Railo, Koivisto, Revonsuo, & Hannula,  
555 2008; Vetter et al., 2008). Other studies point to a crucial role of non-numerical factors  
556 such as visual working memory (Piazza et al., 2011) and stimulus spatial configural  
557 processing (Ashkenazi et al., 2013; Krajcsi, Szabo, & Morocz, 2013; Mandler & Shebo,  
558 1982).

559 For the aims of the present study we reanalysed recent data collected for other purposes.  
560 In a previous study we only considered data in the estimation range, carefully avoiding  
561 subitizing (Anobile et al., 2017), and detected a good correlation between children  
562 simultaneous (spatial arrays) numerosity estimation precision and math. It is important to  
563 note that formal math tasks and scores in the two studies were identical, and therefore the  
564 lack of correlation found here cannot be accounted for by difficulties in measurement of  
565 arithmetic abilities. In any case, since mathematics is not a single concept, the possibility  
566 still remains open that certain skills may be exclusively linked to the ANS and others to the  
567 subitizing system. In previous reports, our group has proposed the ANS and subitizing

568 overlap, but for very low numerosities (subitizing range) precision is boosted by attentional  
569 resources making it particularly fast and precise. Depriving attentional resources, precision  
570 level in the subitizing range approaches that of the estimation range (Anobile et al., 2012;  
571 Burr et al., 2010). Moreover, while in normal conditions subitizing is not susceptible to  
572 numerosity adaptation, it is adaptable when attention is deprived in a dual-task paradigm  
573 (Burr et al., 2011). It would be interesting to test whether performance within the subitizing  
574 range, measured under dual task conditions, correlated with math abilities.

575

### 576 **Subitizing across development**

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

584

### 585 **Subitizing across sensory modalities and stimuli format**

586 The present results confirm previous studies showing that adults can subitize auditory and  
587 visual sequential stimuli (Camos & Tillmann, 2008; Repp, 2007) and go on to show that  
588 this ability is present in primary school children. In children, the capacity to subitize audio  
589 and visual sequential events are positively correlated with each other, indicating a  
590 common system for perception of sequential stimuli. This correlation was strong, with LBF  
591 near 0.5, robust enough to survive Bonferroni correction, and remained significant even  
592 when the important covariates of age and non-verbal reasoning scores (Raven matrices)  
593 were controlled for. With adults we found no significant correlations between subitizing

594 capacities. In light of these results we might hypothesize that sequential and simultaneous  
595 subitizing are subserved by separate mechanisms. As auditory and visual sequential  
596 subitizing capacities are linked in children but not in adults, this may suggest that the  
597 “sequential subitizing system” starts as a cross-sensory system that differentiates later on.  
598 Since in our sample of children we did not find a correlation between simultaneous and  
599 sequential subitizing, this hypothesis predicts that in younger children this correlation  
600 should exist. To indicate which factors may cause the hypothesised differentiation is  
601 difficult to say, but we may speculate that it should reflect a gradually reduced cross-talk  
602 between general domain skills across different sensory modalities (for example auditory  
603 and visual attentional and/or working memory resources). It would be interesting to devise  
604 future studies to test all these hypotheses. Moreover, a lack of correlation cannot be  
605 interpreted as definitive prove of separate mechanisms. Other studies, perhaps using  
606 causative methods, are needed to demonstrate separate mechanisms for simultaneous  
607 and sequential subitizing, as well as to define their relationship with math. As a  
608 complimentary way to test for our current results, it would be interesting to look for  
609 neuropsychological dissociations in patients with brain lesions as the represented data  
610 clearly predict the possibility of deficits selectively affecting simultaneous or sequential  
611 subitizing abilities.

612

## 613 **Conclusions**

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

619

620

621 **References**

- 622 Agrillo, C., Miletto Petrazzini, M. E., & Bisazza, A. (2017). Numerical abilities in fish: A  
623 methodological review. *Behav Processes*. doi:10.1016/j.beproc.2017.02.001
- 624 Anobile, G., Arrighi, R., Castaldi, E., Grassi, E., Pedonese, L., PA, M. M., & Burr, D. C. (2017).  
625 Spatial but Not Temporal Numerosity Thresholds Correlate With Formal Math Skills in  
626 Children. *Dev Psychol*. doi:10.1037/dev0000448
- 627 Anobile, G., Arrighi, R., Togoli, I., & Burr, D. C. (2016). A shared numerical representation for  
628 action and perception. *Elife*, 5. doi:10.7554/eLife.16161
- 629 Anobile, G., Stievano, P., & Burr, D. C. (2013). Visual sustained attention and numerosity  
630 sensitivity correlate with math achievement in children. *J Exp Child Psychol*, 116(2),  
631 380-391. doi:10.1016/j.jecp.2013.06.006
- 632 Anobile, G., Turi, M., Cicchini, G. M., & Burr, D. C. (2012). The effects of cross-sensory  
633 attentional demand on subitizing and on mapping number onto space. *Vision Res*, 74,  
634 102-109. doi:10.1016/j.visres.2012.06.005
- 635 Arp, S., & Fagard, J. (2005). What impairs subitizing in cerebral palsied children?  
636 *Developmental Psychobiology*, 47(1), 89-102. doi:10.1002/dev.20069
- 637 Arp, S., Taranne, P., & Fagard, J. (2006). Global perception of small numerosities (subitizing) in  
638 cerebral-palsied children. *J Clin Exp Neuropsychol*, 28(3), 405-419.  
639 doi:10.1080/13803390590935426
- 640 Arrighi, R., Lunardi, R., & Burr, D. (2011). Vision and audition do not share attentional  
641 resources in sustained tasks. *Front Psychol*, 2, 56. doi:10.3389/fpsyg.2011.00056
- 642 Arrighi, R., Togoli, I., & Burr, D. C. (2014). A generalized sense of number. *Proc Biol Sci*,  
643 281(1797). doi:10.1098/rspb.2014.1791

- 644 Ashkenazi, S., & Henik, A. (2012). Does attentional training improve numerical processing in  
645 developmental dyscalculia? *Neuropsychology*, *26*(1), 45-56. doi:10.1037/a0026209
- 646 Ashkenazi, S., Mark-Zigdon, N., & Henik, A. (2013). Do subitizing deficits in developmental  
647 dyscalculia involve pattern recognition weakness? *Developmental Science*, *16*(1), 35-  
648 46. doi:10.1111/j.1467-7687.2012.01190.x
- 649 Ashkenazi, S., & Henik, A. (2010). Attentional networks in developmental dyscalculia. *Behav*  
650 *Brain Funct*, *6*, 2. doi:10.1186/1744-9081-6-2
- 651 Biancardi, A., Bachmann, C., & Nicoletti, C. (2016). *Batteria per la discalculia evolutiva (BDE2)*.  
652 Trento: centro studi erickson.
- 653 Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis*, *10*(4), 433-436.
- 654 Bruandet, M., Molko, N., Cohen, L., & Dehaene, S. (2004). A cognitive characterization of  
655 dyscalculia in Turner syndrome. *Neuropsychologia*, *42*(3), 288-298.
- 656 Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics  
657 ability: inhibition, switching, and working memory. *Dev Neuropsychol*, *19*(3), 273-293.  
658 doi:10.1207/S15326942DN1903\_3
- 659 Burr, D. C., Anobile, G., & Turi, M. (2011). Adaptation affects both high and low (subitized)  
660 numbers under conditions of high attentional load. *Seeing Perceiving*, *24*(2), 141-150.  
661 doi:10.1163/187847511X570097
- 662 Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires  
663 attentional resources. *J Vis*, *10*(6), 20. doi:10.1167/10.6.20
- 664 Camos, V., & Tillmann, B. (2008). Discontinuity in the enumeration of sequentially presented  
665 auditory and visual stimuli. *Cognition*, *107*(3), 1135-1143.  
666 doi:10.1016/j.cognition.2007.11.002
- 667 Carey, S. (2002). Cognitive foundations of arithmetic: evolution and ontogenesis. *Mind and*  
668 *Language*, *16*, 37-55 doi:10.1111/1468-0017.00155

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

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

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

676 Cowan, N., Morey, C. C., AuBuchon, A. M., Zwilling, C. E., & Gilchrist, A. L. (2010). Seven-year-  
677 olds allocate attention like adults unless working memory is overloaded. *Dev Sci,*  
678 *13*(1), 120-133. doi:10.1111/j.1467-7687.2009.00864.x

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

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

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

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

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

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

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

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

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

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

704 Gray, S. A., & Reeve, R. A. (2014). Preschoolers' Dot Enumeration Abilities Are Markers of  
705 Their Arithmetic Competence. *PLoS One*, *9*(4). doi:ARTN e94428  
706 10.1371/journal.pone.0094428

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

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

712 Kaufman, E. L., & Lord, M. W. (1949). The discrimination of visual number. *Am J Psychol*,  
713 *62*(4), 498-525.

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

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

718 Mandler, G., & Shebo, B. J. (1982). Subitizing - an Analysis of Its Component Processes. *Journal*  
719 *of Experimental Psychology-General*, *111*(1), 1-22. doi:Doi 10.1037/0096-  
720 3445.111.1.1

721 McLachlan, N. M., Marco, D. J. T., & Wilson, S. J. (2012). Pitch Enumeration: Failure to Subitize  
722 in Audition. *PLoS One*, *7*(4). doi:ARTN e33661  
723 10.1371/journal.pone.0033661

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

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

729 Olivers, C. N. L., & Watson, D. G. (2008). Subitizing requires attention. *Visual Cognition*, *16*(4),  
730 439-462. doi:10.1080/13506280701825861

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

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

737 Paterson, S. J., Girelli, L., Butterworth, B., & Karmiloff-Smith, A. (2006). Are numerical  
738 impairments syndrome specific? Evidence from Williams syndrome and Down's  
739 syndrome. *J Child Psychol Psychiatry*, *47*(2), 190-204. doi:10.1111/j.1469-  
740 7610.2005.01460.x

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

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

746 Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., . . . Zorzi, M.  
747 (2010). Developmental trajectory of number acuity reveals a severe impairment in  
748 developmental dyscalculia. *Cognition*, *116*(1), 33-41.  
749 doi:10.1016/j.cognition.2010.03.012

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

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

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

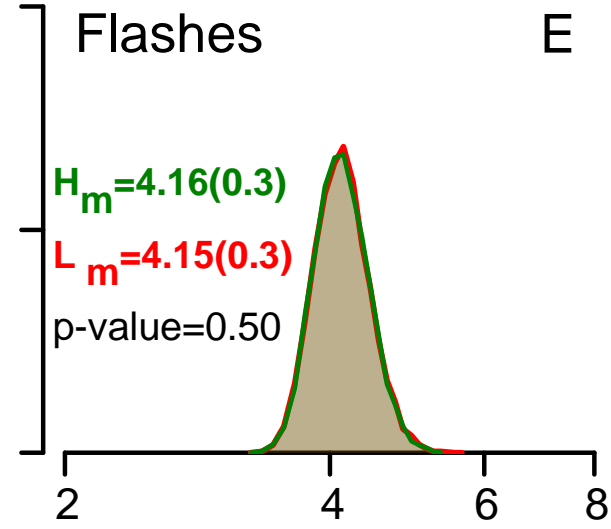
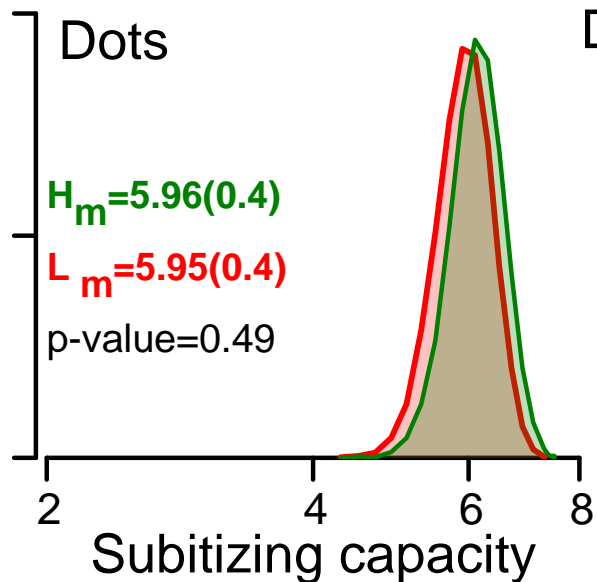
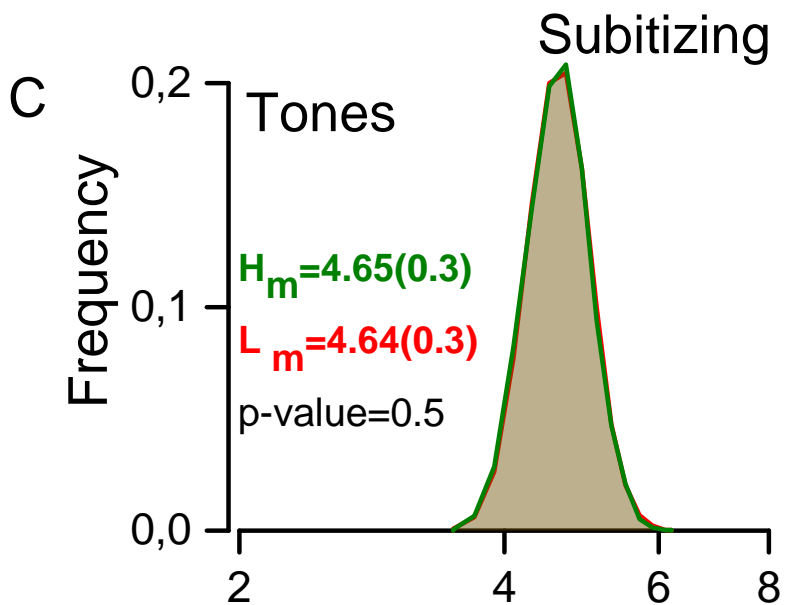
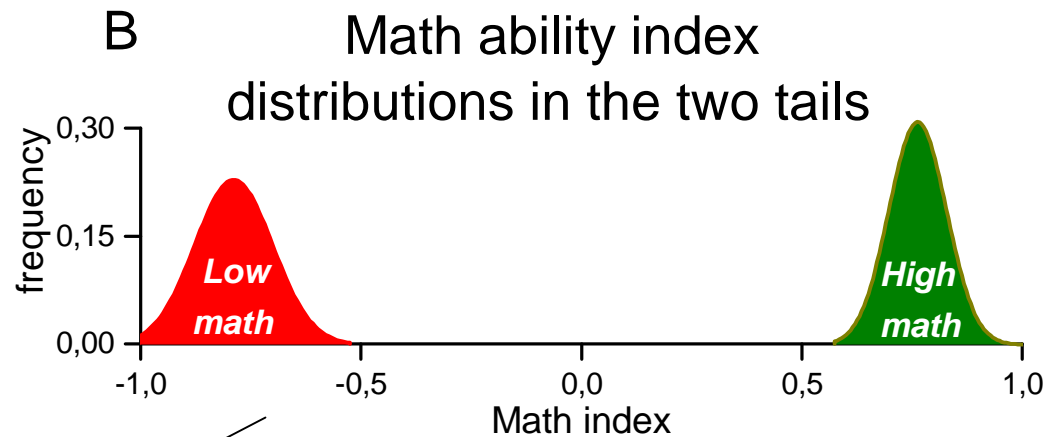
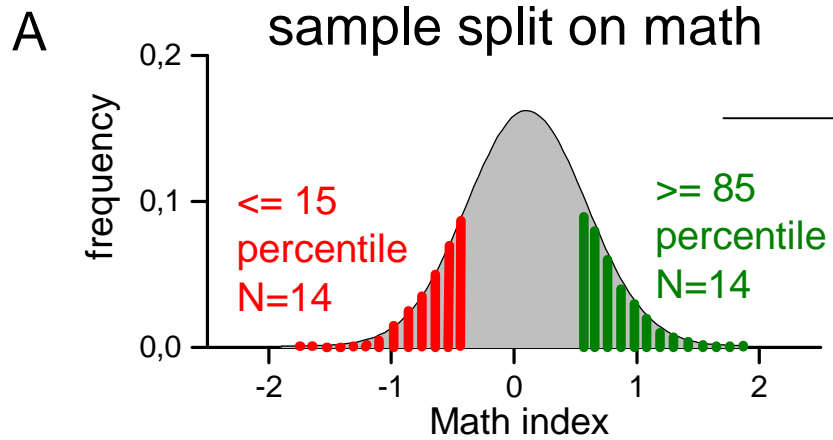
758 Plaisier, M. A., Tiest, W. M. B., & Kappers, A. M. L. (2010). Grabbing subitizing with both hands:  
759 bimanual number processing. *Experimental Brain Research*, *202*(2), 507-512.  
760 doi:10.1007/s00221-009-2146-1

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

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

- 766 Railo, H., Koivisto, M., Revonsuo, A., & Hannula, M. M. (2008). The role of attention in  
767 subitizing. *Cognition*, *107*(1), 82-104. doi:10.1016/j.cognition.2007.08.004
- 768 Reeve, R., Reynolds, F., Humberstone, J., & Butterworth, B. (2012). Stability and change in  
769 markers of core numerical competencies. *J Exp Psychol Gen*, *141*(4), 649-666.  
770 doi:10.1037/a0027520
- 771 Repp, B. H. (2007). Perceiving the numerosity of rapidly occurring auditory events in metrical  
772 and nonmetrical contexts. *Percept Psychophys*, *69*(4), 529-543.
- 773 Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect  
774 numerical estimation? *Psychological Science*, *19*(6), 607-614. doi:Doi 10.1111/J.1467-  
775 9280.2008.02130.X
- 776 Riggs, K. J., Ferrand, L., Lancelin, D., Fryziel, L., Dumur, G., & Simpson, A. (2006). Subitizing in  
777 tactile perception. *Psychological Science*, *17*(4), 271-272. doi:10.1111/j.1467-  
778 9280.2006.01696.x
- 779 Rugani, R., Vallortigara, G., Priftis, K., & Regolin, L. (2015). Animal cognition. Number-space  
780 mapping in the newborn chick resembles humans' mental number line. *Science*,  
781 *347*(6221), 534-536. doi:10.1126/science.aaa1379
- 782 Schleifer, P., & Landerl, K. (2011). Subitizing and counting in typical and atypical development.  
783 *Dev Sci*, *14*(2), 280-291.
- 784 Steele, A., Karmiloff-Smith, A., Cornish, K., & Scerif, G. (2012). The multiple subfunctions of  
785 attention: differential developmental gateways to literacy and numeracy. *Child Dev*,  
786 *83*(6), 2028-2041. doi:10.1111/j.1467-8624.2012.01809.x
- 787 Toll, S. W., Kroesbergen, E. H., & Van Luit, J. E. (2016). Visual working memory and number  
788 sense: Testing the double deficit hypothesis in mathematics. *Br J Educ Psychol*, *86*(3),  
789 429-445. doi:10.1111/bjep.12116

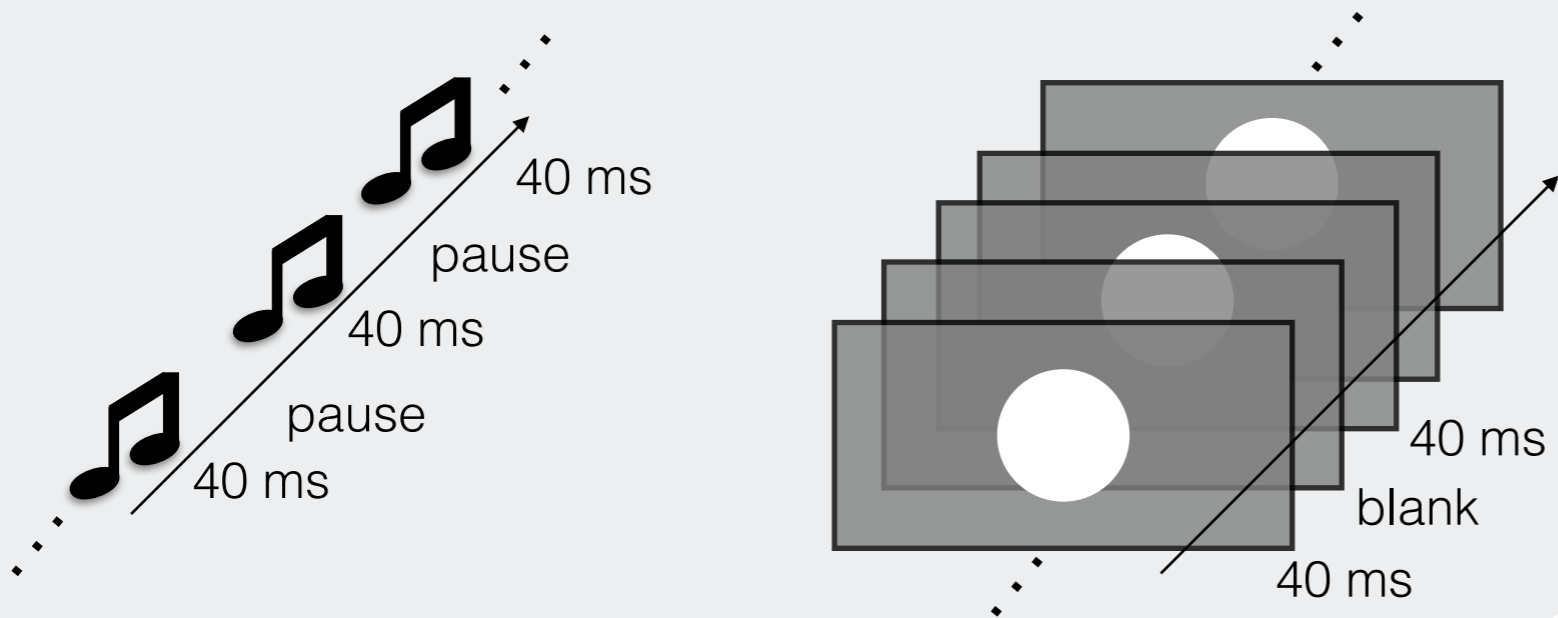
- 790 Vetter, P., Butterworth, B., & Bahrami, B. (2008). Modulating attentional load affects  
791 numerosity estimation: evidence against a pre-attentive subitizing mechanism. *PLoS*  
792 *One*, 3(9), e3269. doi:10.1371/journal.pone.0003269
- 793 Wang, J. J., Halberda, J., & Feigenson, L. (2017). Approximate number sense correlates with  
794 math performance in gifted adolescents. *Acta Psychol (Amst)*, 176, 78-84.  
795 doi:10.1016/j.actpsy.2017.03.014
- 796 Wetzels, R., & Wagenmakers, E. J. (2012). A default Bayesian hypothesis test for correlations  
797 and partial correlations. *Psychon Bull Rev*, 19(6), 1057-1064. doi:10.3758/s13423-  
798 012-0295-x
- 799



# Numerosity tasks: *how many?*

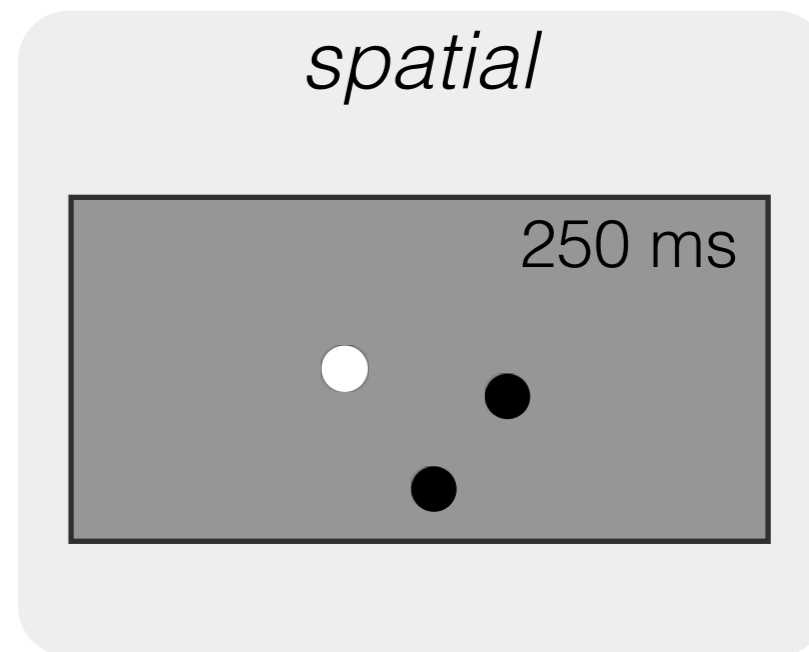
A

*temporal*



B

*spatial*

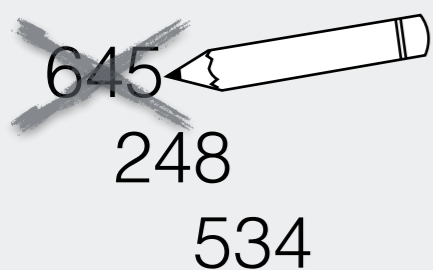


C

# Arithmetic tasks

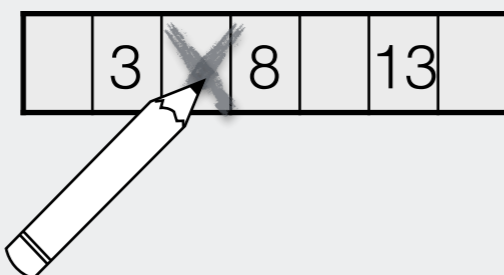
*semantic tasks*

'cross out the largest'



'numberline'

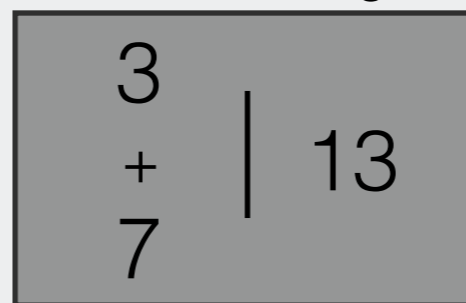
place number: 5



D

*mental addition*

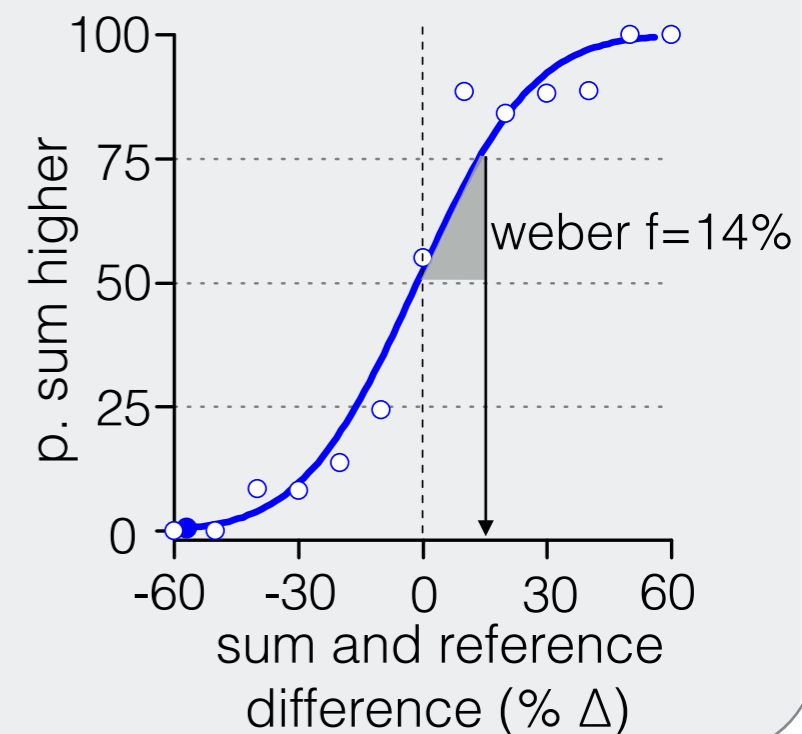
Which side is higher?

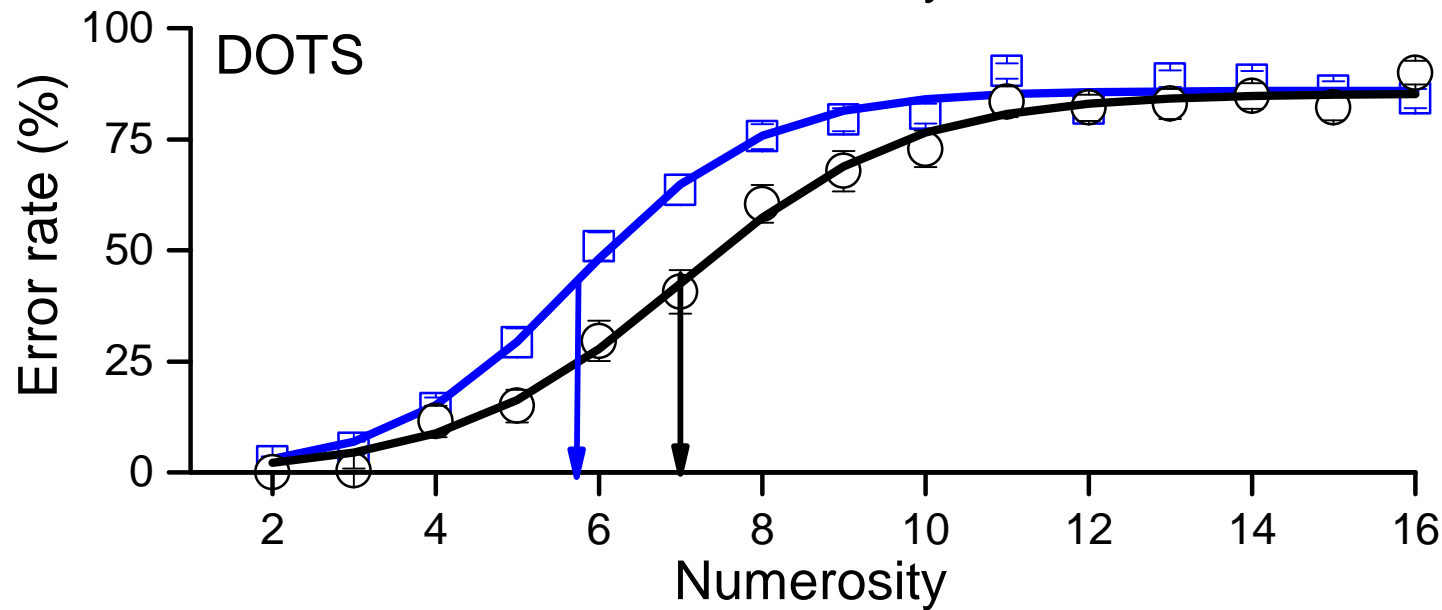
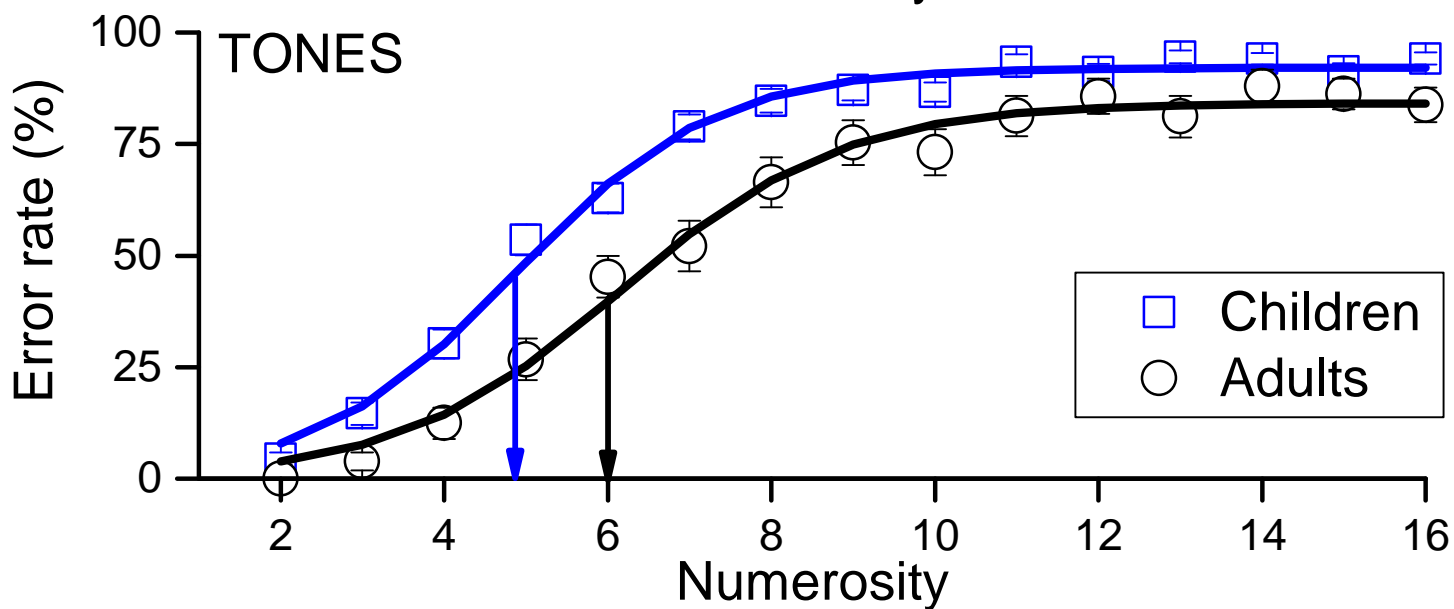
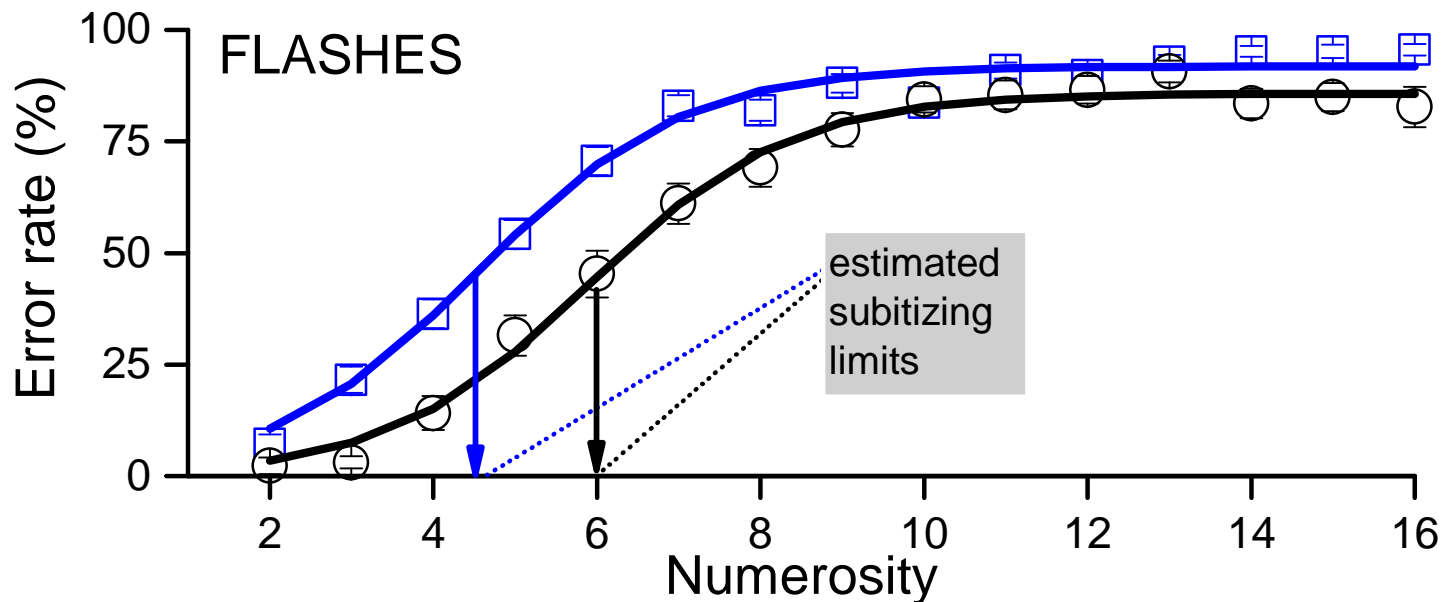


*sum*

*reference*

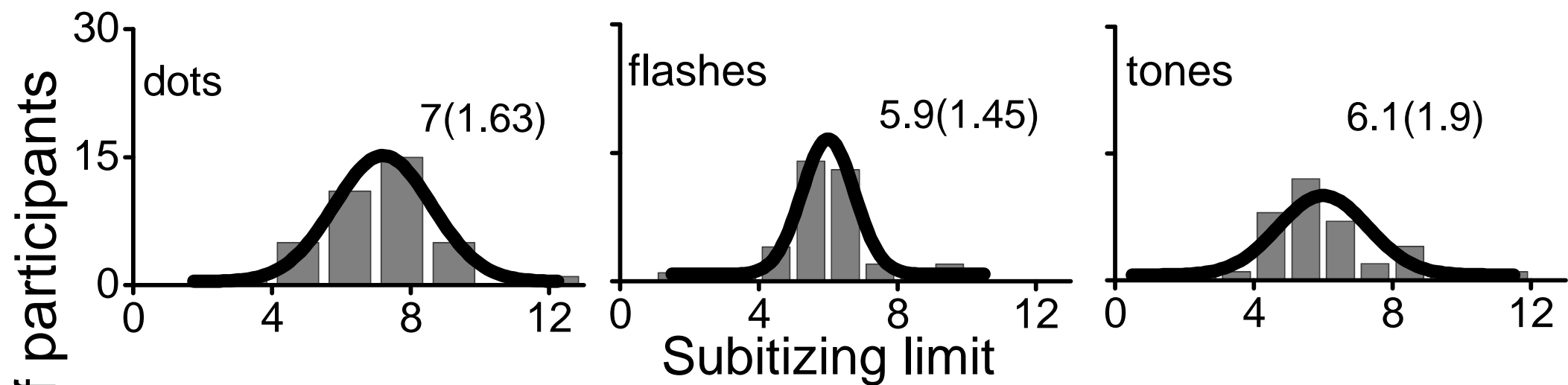
e.g.  $\Delta +30\%$





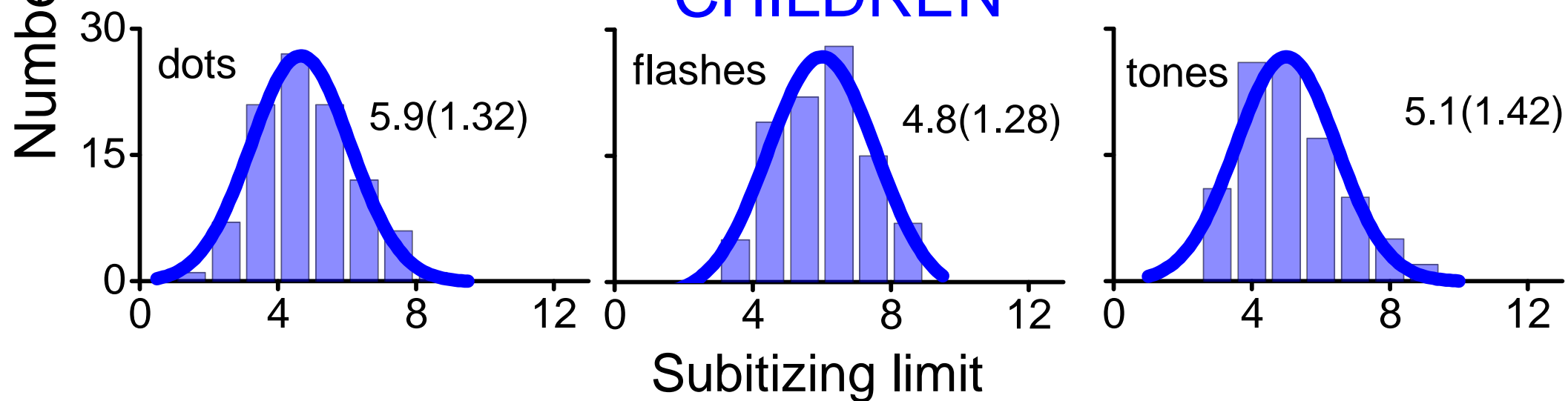
# ADULTS

A

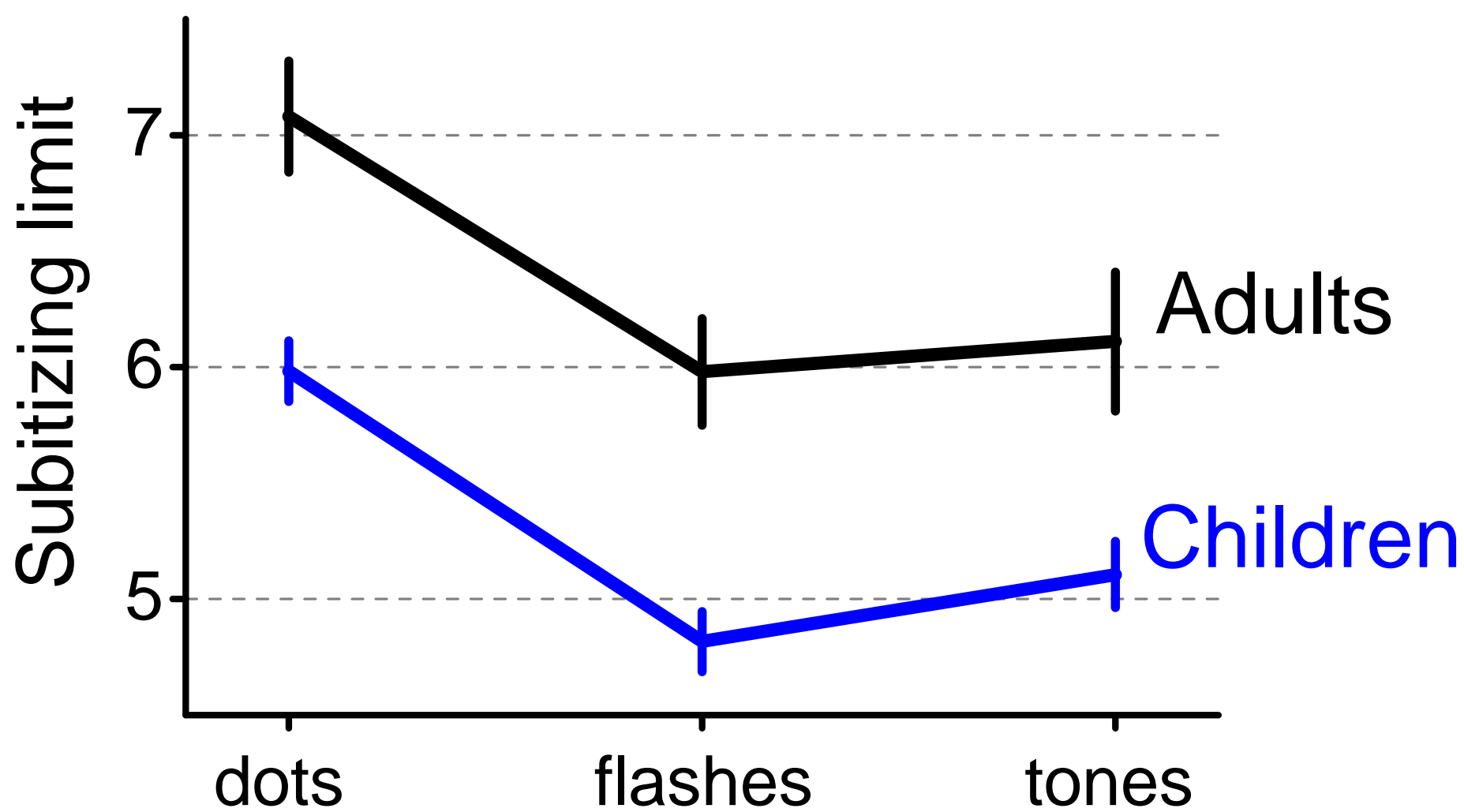


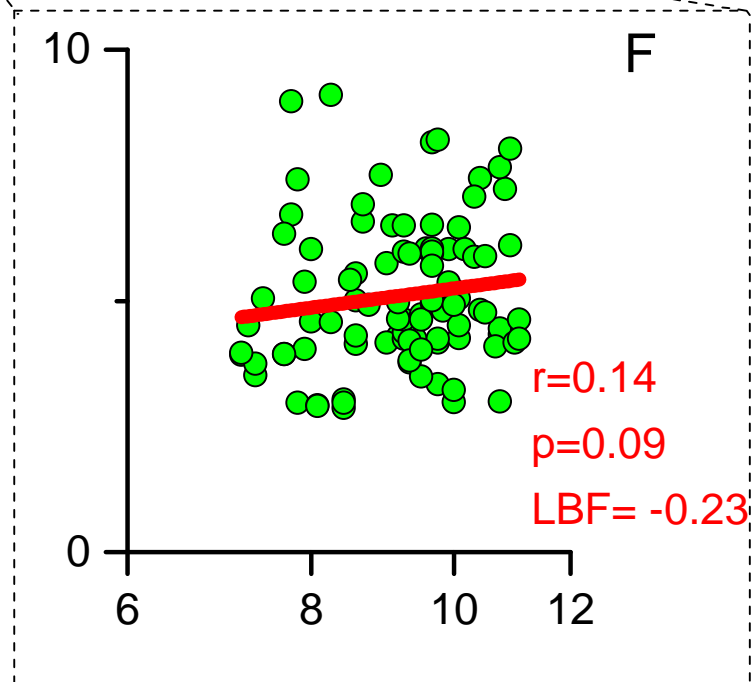
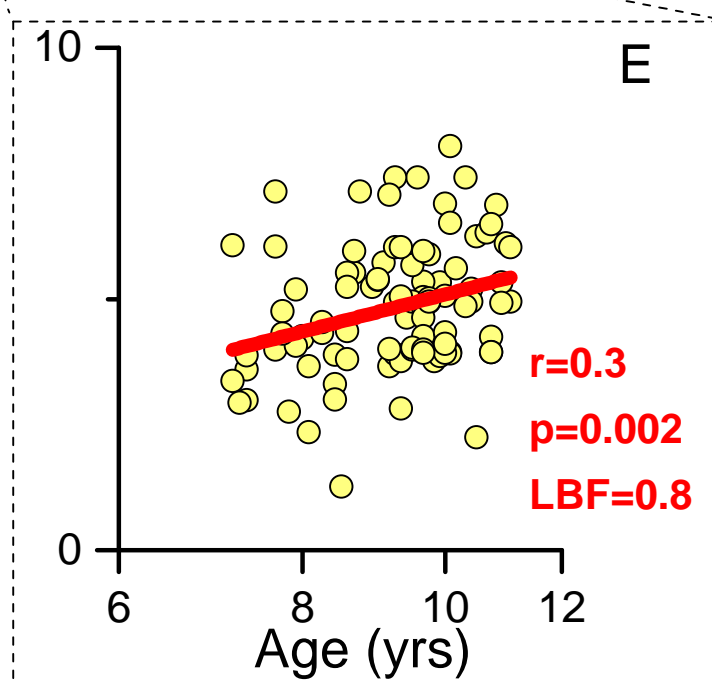
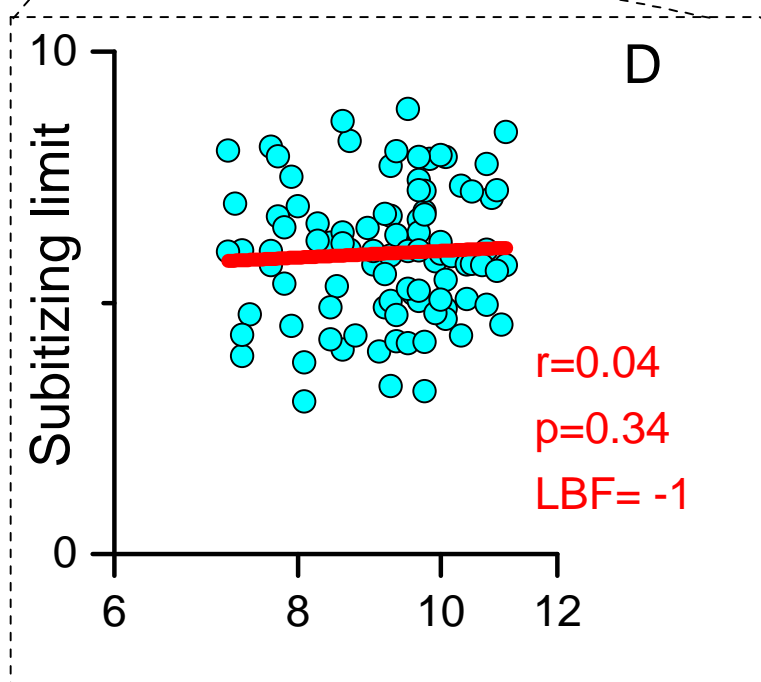
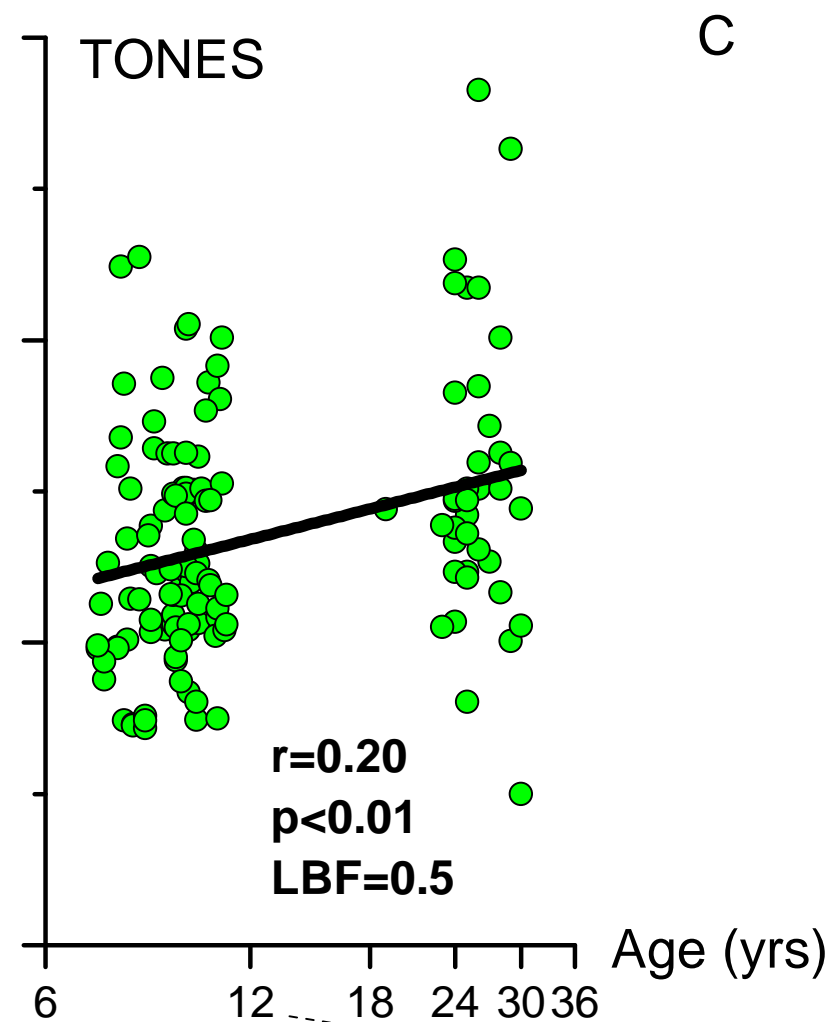
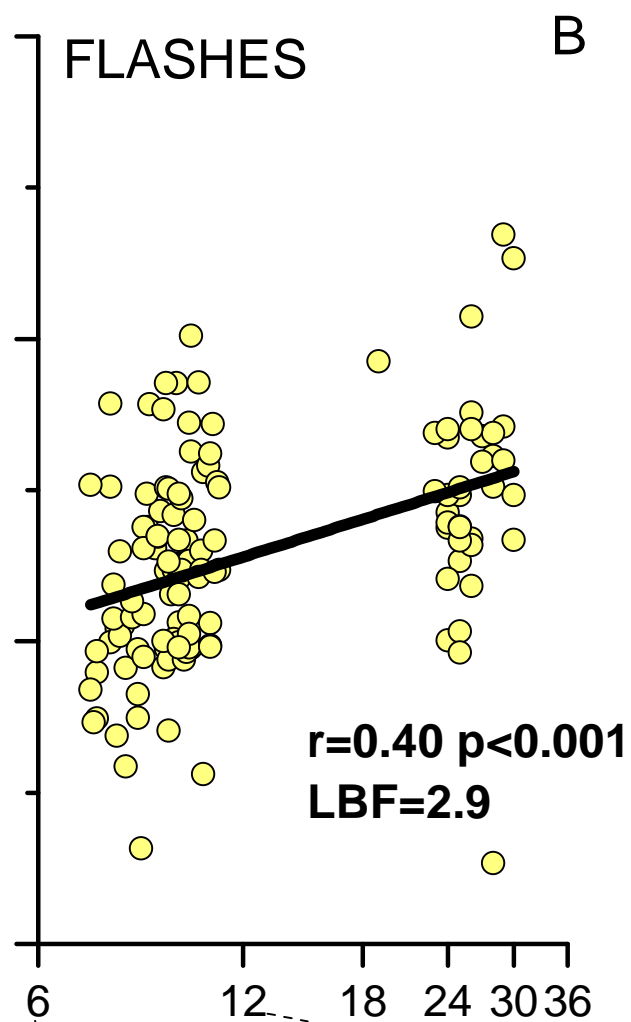
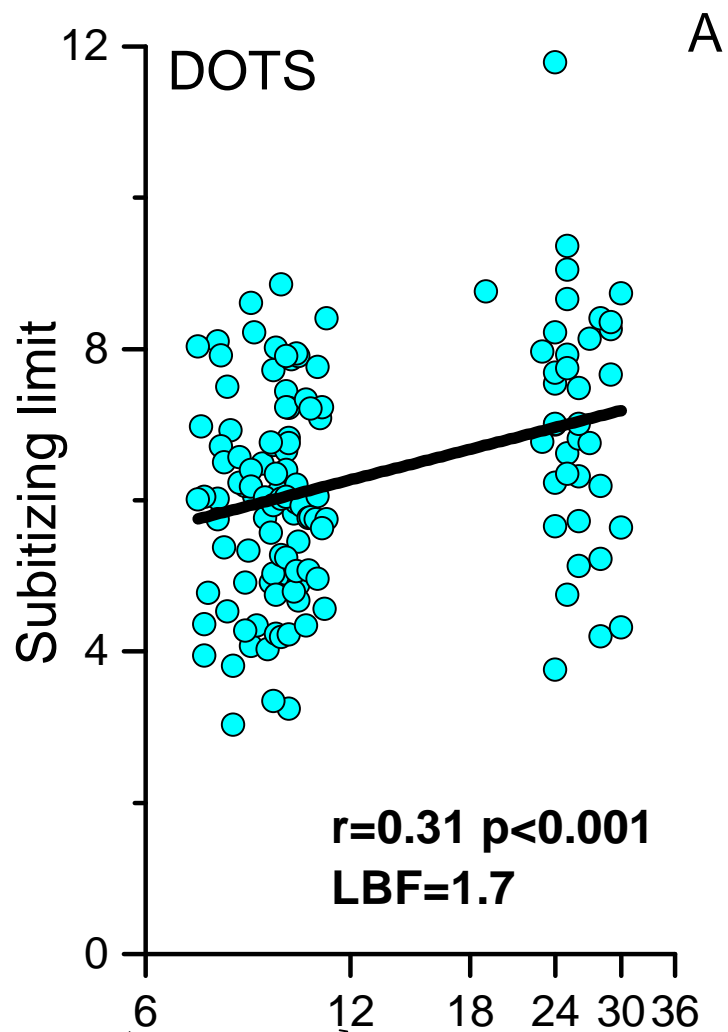
# CHILDREN

B

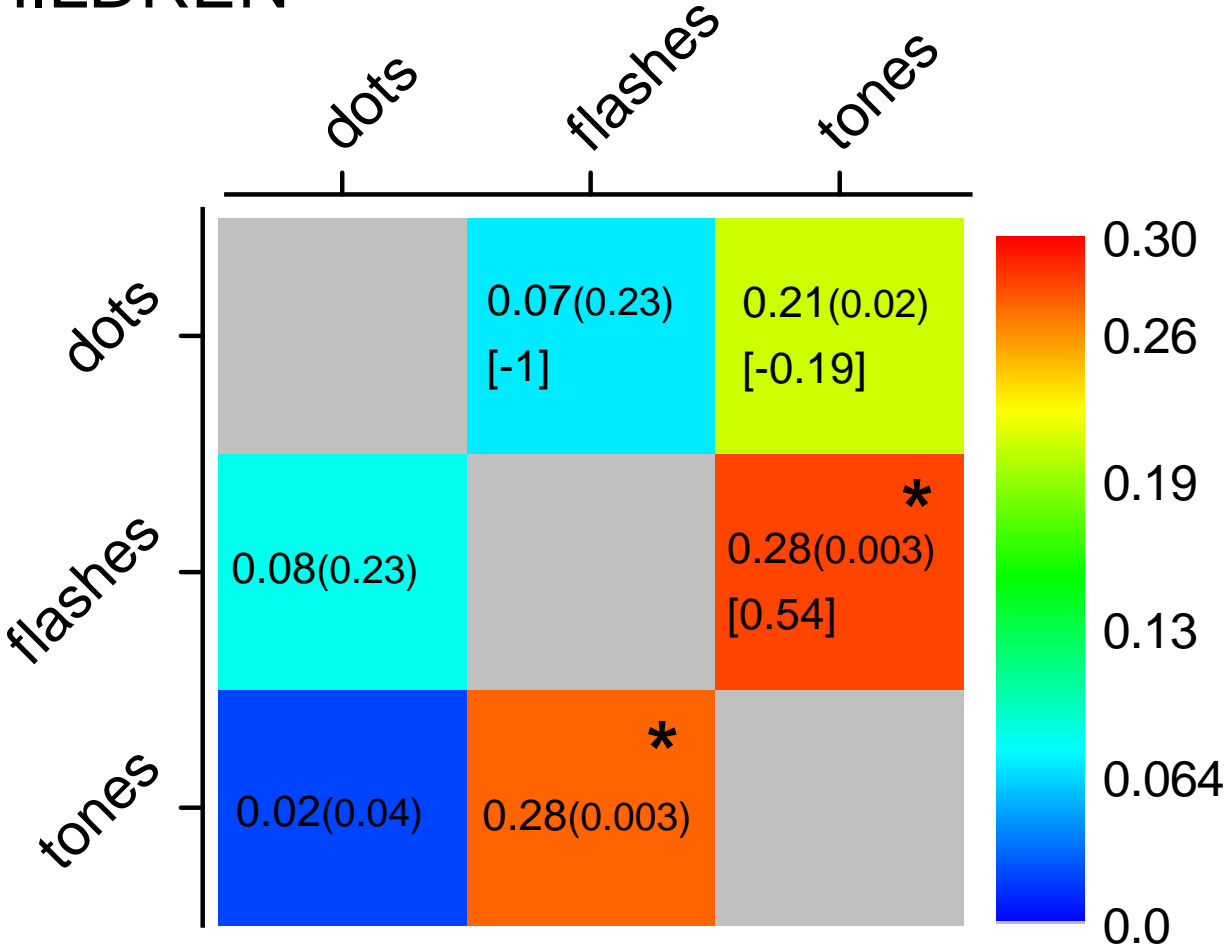


C





# CHILDREN

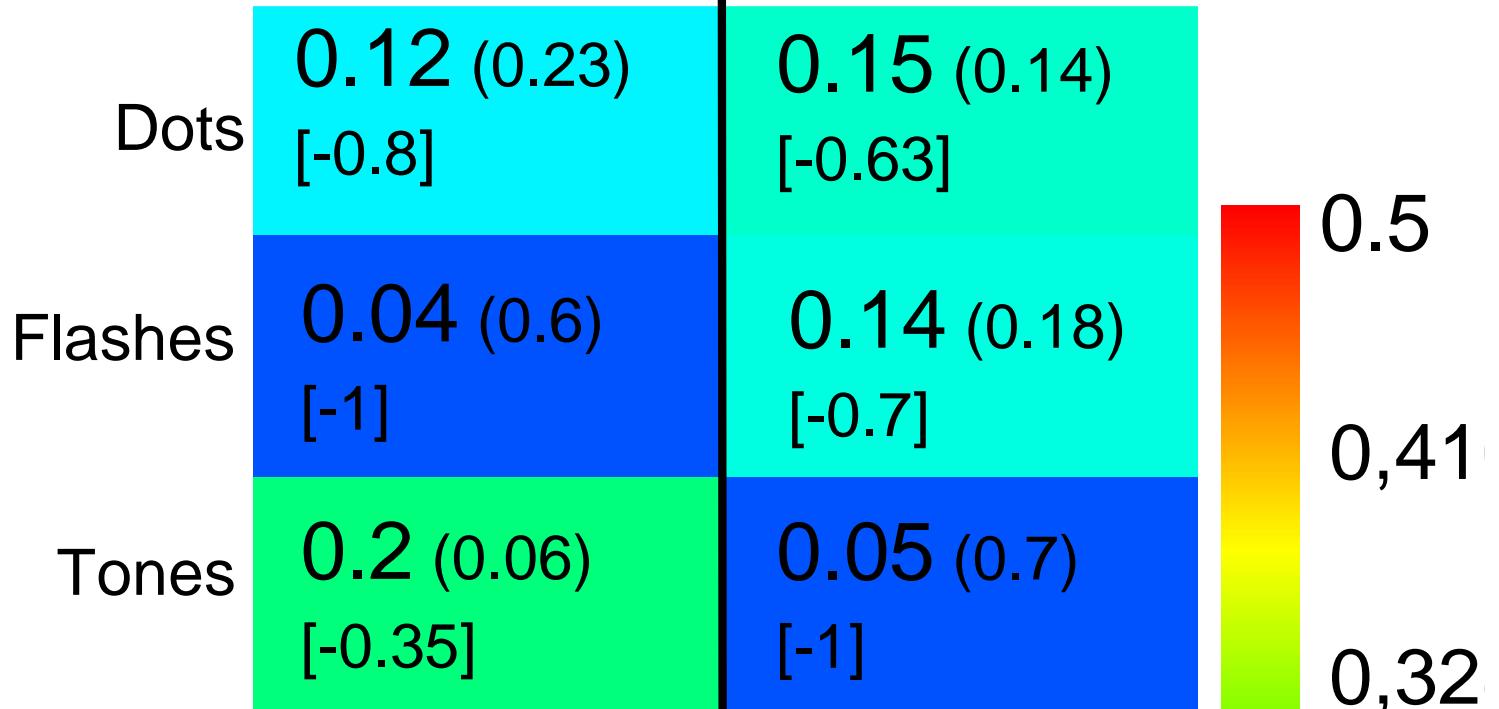


Bonferroni  $\alpha=0.005$  (0.05/9)

Mental  
addition

Semantic  
tasks

*Children*



*Adults*

