Journal of Experimental Child Psychology 178 (2019) 86–103

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

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

Giovanni Anobile [a,*], Roberto Arrighi [b], David C. Burr [b,c,d]

[a] Department of Developmental Neuroscience, Stella Maris Scientific Institute, 56018 Calambrone (Pisa), Italy
[b] Department of Neuroscience, Psychology, Pharmacology and Child Health, University of Florence, 50139 Florence, Italy
[c] School of Psychology, University of Sydney, Sydney, NSW 2006, Australia
[d] Department of Translational Research on New Technologies in Medicines and Surgery, University of Pisa, 56123 Pisa, Italy

ARTICLE INFO

Article history:
Received 20 February 2018
Revised 24 September 2018
Available online 26 October 2018

Keywords:
Numerosity
Subitizing
Numerical cognition
Approximate number system
Dyscalculia
Cross-modal perception

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 addressed two issues: (a) whether subitizing generalizes over modalities and stimulus formats and (b) whether subitizing correlates 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) or simultaneously (visual presentations, dot arrays). The results show that (a) subitizing limits for adults were one item larger than those 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 subitizing of sequences of sounds are strongly correlated with each other in children; and (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 subserved by separate systems. Furthermore, subitizing does not seem to be related to numerical abilities.

* Corresponding author.
E-mail address: g.anobile@fsm.unipi.it (G. Anobile).

https://doi.org/10.1016/j.jecp.2018.09.017
0022-0965/

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 for up to 4–6 items; after this range, performance decreases monotonically—both reaction times (RTs) and accuracy. Kaufman and Lord (1949) were first to coin the term "subitizing," derived from the Latin *subitus,* meaning sudden. Subitizing can be defined in several ways. Kaufman and Lord (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 RTs (Piazza et al., 2011; Revkin et al., 2008). This method tends to overestimate the limit compared with other techniques, but it has been proven to be robust and well suited to detect interindividual differences.

Subitizing phenomena are thought to be linked to many non-numerical capacities such as 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 1 year of age (Piazza, 2010), whereas spatial working memory continues to develop until 6 or 7 years of age (Cowan, Morey, AuBuchon, Zwilling, & Gilchrist, 2010). In addition, OTS capacity measured by a visual multiple object tracking task is affected by visual, but not auditory, attentional deprivation (Arrighi et al., 2011), whereas 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 the OTS (Anobile, Stievano, & Burr, 2013) and math performance. In brief, although subitizing has been extensively studied, the underlying mechanisms and how it relates to other cognitive capacities are still unclear.

Moreover, although 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), whereas more evidence has been provided for the existence of subitizing sequences of haptic stimuli (Ferrand, Riggs, & Castronovo, 2010; Gallace, Tan, Haggard, & Spence, 2008; Plaisier & Smeets, 2011; Plaisier, Bergmann Tiest, & Kappers, 2009; 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 the 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 participants do not show clear peculiarities for either the OTS or subitizing (Piazza, 2010). One study involving dyscalculic adolescents reported almost typical subitizing capacity (Ceulemans et al., 2014), whereas another reported that 43–79% of dyscalculic participants in the age range of 7–17 years had impaired subitizing, more evident in older participants (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's 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 was to investigate subitizing, testing for generalization across format and modality. We studied its developmental trend and looked 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.

Method

Participants

A total of 98 children (7.1–11.0 years old, mean = 9.2 years) and 38 adults (19–30 years, mean = 25.6 years) 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-in. LG touchscreen monitor with 1280 $\times$ 1024 resolution at a refresh rate of 60 Hz. Each participant was tested in two separate sessions (usually within the same week) lasting about 1 h 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 nonverbal reasoning task (Raven matrices). Math skills and nonverbal reasoning were measured at the end of the first session, and perceptual tasks were administered in a pseudorandom order between participants. This study was based on a new analysis of a set of data collected for other purposes (Anobile et al., 2018). The experimental methods used here were the same as those used in the previous study, but here we focused on the subitizing range by not excluding numerosities $\leq$4 from the analyses.

Numerosity estimation

Visual stimuli were either ensembles of 0.5° diameter dots half-white and half-black in order to balance luminance across numerosities (in the 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 disks of 90 cd m^2 and 5° diameter), presented in a pseudorandom order within a 2-s interval (Fig. 1A 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 the participant viewing distance set at 57 cm, on a gray 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 participants 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' judgments (mainly those of young children) to have all estimates within the numerical range without aberrant responses (for a similar procedure, see Revkin et al., 2008). After training was completed, the testing phase started with a block of 51 trials (three repetitions for each numerosity),

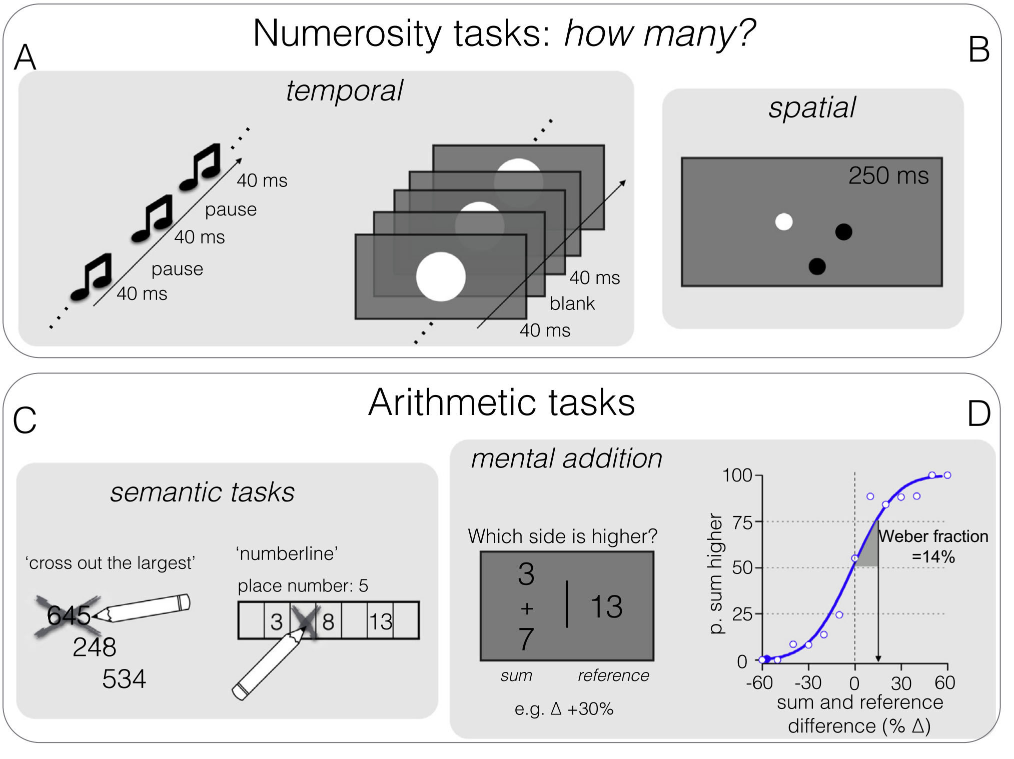

Fig. 1. Tasks and paradigms. (A, B) Each trial started with a fixation point (lasting until the experimenter pressed the space bar), followed by either a series of beeps or white disk flashes (A) or a cloud of dots simultaneously presented (B). Participants verbally reported perceived numerosity. (C) Children were asked to solve a series of tasks where they needed to recognize and cross out the numerically larger digit among three digits 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 digit numbers on the left and to compare the sum with that on the right (13 in this example) and indicate which side was numerically higher (right in this example). Weber fraction measured precision: In the sample psychometric function reported, a Weber's fraction of 0.14 indicates that the sum of the two addenda needed to be 14% higher or lower than reference to raise responses from chance (50%) to 75% correct responses. Stimuli remained until response.

with no feedback. Each participant performed 204 trials in total. Test numerosity ranged from 2 to 18, but we computed error rates and fitted (see later) only the range from 2 to 16 to avoid edge effects. Average temporal rates for both flash and sound stimuli were 640 ms (range = 110–1180), 500 ms (range = 130–900), and 400 ms (range = 140–650) for numerosities 2, 3, and 4, respectively. Because the counting speed for numbers in Italian primary school children is about 600–800 ms per number, and the stimulus sequences were not regular but rather jittered in time, it is unlikely that children were able to serially count the stimuli.

Semantic skills

Two paper-and-pencil tasks were administered (Fig. 1C), asking participants to (a) mark the largest number in a set of three (one to five digits, 36 trials) and (b) 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 but 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 predictors of children's numerosity discrimination thresholds (Anobile et al., 2013, 2018; 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 used by Anobile et al., 2018).

Computerized mental addition task

On each trial, three digits ($3^\circ \times 3^\circ$, 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 to compare the result with the single digit on the right (Fig. 1D). 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 into two separate blocks of 35 each. We applied a time threshold (2 and 5 s for adults and children, respectively), with thresholds being derived from pilot data. In trials where RTs exceeded the thresholds of 5.6% and 1.8% for children and adults, respectively, we gave auditory feedback. The feedback did not provide any information about the accuracy, only about the need to perform the operations more quickly.

Not every trial where RTs exceeded the threshold were eliminated from the analysis because we applied a within-participant cutoff; for each participant, we measured the average RTs (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 function. The percentage difference needed to move from 50% to 75% correct responses, providing a "mental additional discrimination threshold." This is logically equivalent to the Weber fraction usually measured for numerosity discrimination tasks, and it could be interpreted as the amount of noise present in the mental addition process (see Fig. 1D). Similarly to previous studies (Anobile, Castaldi, Turi, Tinelli, & Burr, 2016; Anobile et al., 2018), we computed for each participant two separate *z* scores: one for precision (Weber fraction) and the other for response speed (RT). The *z* scores were measured using the mean and standard deviation of the participant grade class (from second grade 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 numerosities above the subitizing range (Anobile et al., 2018).

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 Fig. 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 correlation procedures. Statistical significance was indexed by *p* values and Bayes factor (Wetzels & Wagenmakers, 2012). The Bayes factor is the ratio of the likelihood probabilities of the two models, with a correlation quantifying 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 is the likelihood that the correlation does not exist. By convention, a log Bayes factor (LBF) greater than 0.5 is considered substantial evidence in favor of the existence of the correlation, and an LBF less than −0.5 is considered substantial evidence in favor 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 the pairwise deletion method.

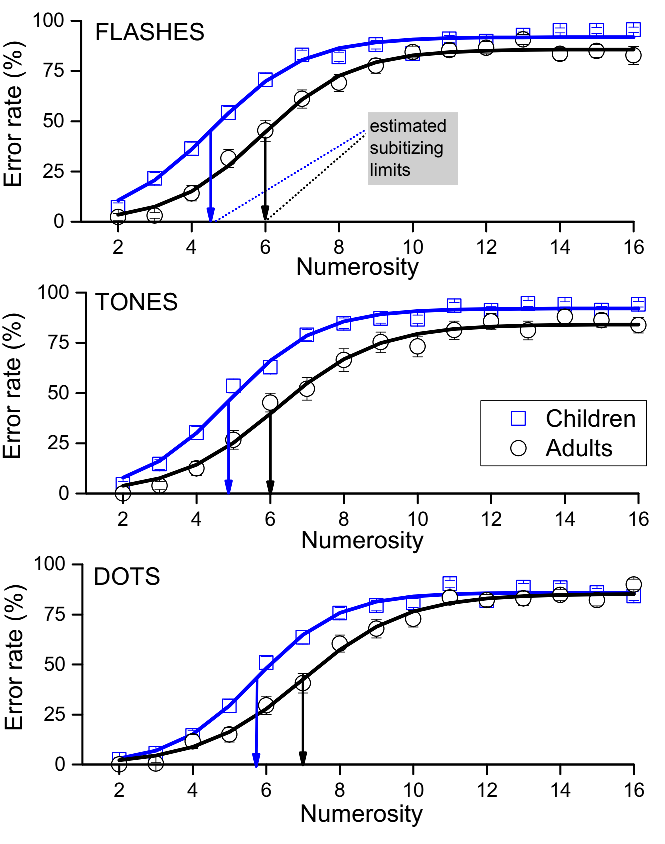

Fig. 2. Error rates as a function of numerosity. Panels from top to bottom report performance for numerosity estimation of flashes, tones, and arrays of dots averaged across participants. Data in blue (upper lines, squares) refer to children, and data in black (lower lines, circles) refer to adults. Error bars are standard errors of the means. Lines are sigmoid functions. Arrows indicate subitizing capacities measured from the inflection point of the fitting functions.

Task reliability

We measured reliability using split-half bootstrap techniques.

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 estimates of reliability. This method was described and validated by Anobile et al. (2018).

Subitizing limits

The R^2 values of the fits were reasonably high, suggesting that 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 values lower then .25 (10% overall). In the second analysis, we looked at pooled data. First, for each numerosity, we pooled together all the trials. Second, separately for each numerosity, we divided the trials into two equal-*n* samples by randomly sampling the data (half the size of the data set taken, sampled with replacement from the full data set). Third, we fitted these two separate data samples with the procedure described above, producing two measurements of subitizing limits. On each reiteration (1000 times), we calculated the difference in the limits of the two conditions and determined 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.

Results

Task reliability

Table 1 reports split-half reliability levels (Pearson's *r*) for all the tasks for participant-by-participant analyses. Indexes all were reasonably good, ranging from .57 to .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 and adults.

Task	Pearson's *r*
Dots	C: 0.64 ± 0.18 A: 0.65 ± 0.19
Flashes	C: 0.57 ± 0.24 A: 0.64 ± 0.20
Tones	C: 0.68 ± 0.13 A: 0.74 ± 0.17
Mental addition	Precision (weber fraction) C: 0.58 ± 0.18 A: 0.75 ± 0.08 Speed (reaction time) C: 0.97 ± 0.006 A: 0.95 ± 0.01

Note. Errors reflect standard errors. C, children; A, adults.

Table 2
Split-half averages and subitizing capacities for children and adults.

Stimuli	Subitizing capacities		Average difference
	First half	Second half	First half versus 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 = .98) A: 0.012 ± 0.29 (p = .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 = .93) A: 0.006 ± 0.24 (p = .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 = .99) A: 0.018 ± 0.32 (p = .96)

Note. Analyses were performed on pooled data. Errors reflect standard errors of the mean. C, children; A, adults.

Goodness of fit

The sigmoid fits describe the data well (Fig. 1), with good coefficients of determination (R^2). Indeed, between-participants average R^2 values for children were .74 (SD = .14, min = .32), .73 (SD = .17, min = .33), and .76 (SD = .37, min = .37), for dots, flashes, and tones estimations, respectively. R^2 fits for adults were on average .79 (SD = .14, min = .25), .76 (SD = .17, min = .42), and .75 (SD = .18, min = .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. Similarly to Piazza et al. (2010), we adopted a criterion of eliminating participants with $R^2 < .25$. In total, 10 children had 12 low R^2 fit values (2 fits for dot stimuli, 6 for flashes, and 4 for sounds), and 2 adults had poor R^2 values (1 for flash stimuli and 1 for sounds).

Subitizing limits in children and adults

We measured error rates (symbols in Fig. 2) for estimating numerosity of dot arrays, sequences of flashes, and auditory (tone) events. We fitted errors with sigmoid functions and took the inflection point as an index of subitizing limit. Fig. 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, such as simultaneous spatial ensembles, can be subitized during early childhood.

The same fitting procedure was applied separately to each participant. Fig. 3A and B show the frequency distributions of subitizing limits across participants. On average, limits peaked at about 5 or 6 items (overall averages, pooling together data for all stimuli, were 5.26 and 6.30 for children and adults, respectively), a range often reported in the literature (for a similar value in the case of dots, see Kaufman & Lord, 1949). More important, the distributions show much interparticipant 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).

To monitor developmental changes, we first computed average subitizing as a function of stimulus condition separately for adults and children (Fig. 3C). From inspection, it is clear that the adult limits of subitizing were roughly one element higher than those for primary school children. To statistically test the difference between children and adults, we performed a 2 (Group: children or adults) × 3 (Stimuli: flashes, dots, or sounds) analysis of variance (ANOVA) with subitizing limits as the dependent variable. The analyses confirmed that adults had higher subitizing limits, $F(1, 397) = 46$, $p < .001$, $\eta^2 = .097$. The effect of stimuli was also significant, $F(2, 397) = 19$, $p < .001$, $\eta^2 = .075$, with no interaction with group, $F(2, 397) = 0.085$, $p < .91$, $\eta^2 = .000$, 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, whereas visual and auditory temporal limits were very similar to each other (Table 3).

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

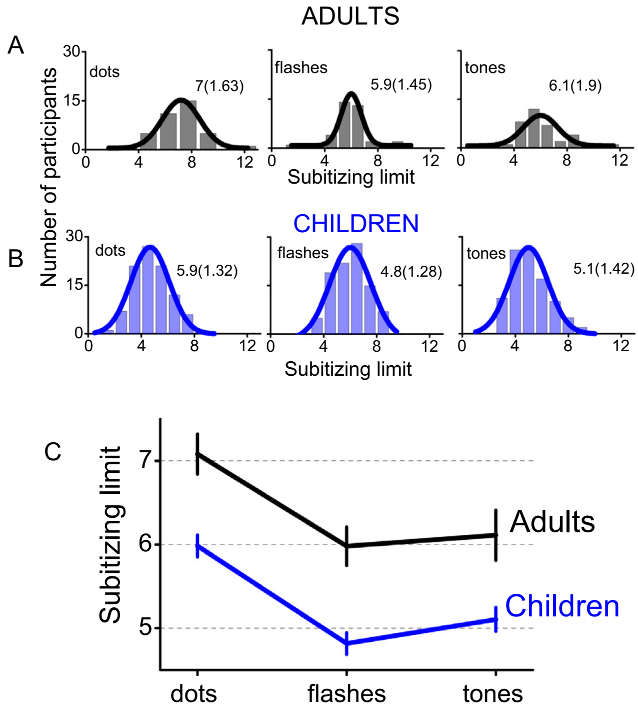

Fig. 3. Frequency distributions of subitizing capacities. (A, B) Panels from left to right report data for subitizing of different stimuli: simultaneous numerosity and sequential numerosity for visual and auditory stimuli, respectively. The first row (A) shows data for adults (in black), and the second row (B) shows data for children (in blue). (C) Average subitizing capacity as a function of stimuli for children (blue, bottom line) and adults (black, top line). Error bars show standard errors of the means. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Difference between averaged subitizing capacities across stimuli for children and adults.

Stimuli	Difference	95% confidence interval of the difference	
		Low	High
Dots versus flashes	C: 1.16*** A: 1.09**	C: 0.69 A: 0.16	C: 1.63 A: 2.03
Dots versus tones	C: 0.87*** A: 0.97*	C: 0.41 A: 0.04	C: 1.34 A: 1.90
Tones versus flashes	C: 0.28 (*ns*) A: 0.13 (*ns*)	C: −0.75 A: −0.80	C: 0.17 A: 1.06

Note. C, children; A, adults.
* $p < .05$ (two-tailed, t test).
** $p < .01$ (two-tailed, t test).
*** $p < .001$ (two-tailed, t test).

childhood to adulthood (Fig. 4A–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-tailed p values; dots: $r = .04$, $p = .34$, LBF = −1; flashes: $r = .29$, $p = .002$, LBF = 0.8; tones: $r = .14$,

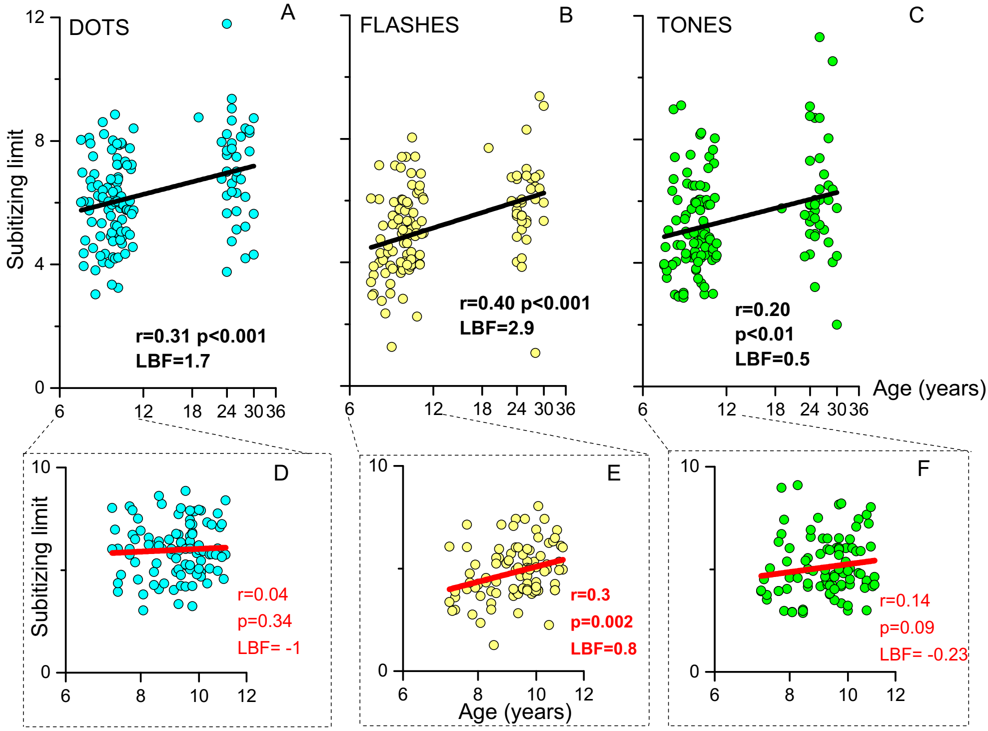

Fig. 4. Developmental trajectories. Panels from left to right report data for subitizing capacities as a function of participant's age for children and adults: (A, D) dots; (B, E) flashes; (C, F) tones.

Table 4
Correlations between subitizing limits.

CHILDREN

	dots	flashes	tones
dots		0.07(0.23) [-1]	0.21(0.02) [-0.19]
flashes	0.08(0.23)		* 0.28(0.003) [0.54]
tones	0.02(0.04)	* 0.28(0.003)	

Colour scale: 0.30, 0.26, 0.19, 0.13, 0.064, 0.0

Bonferroni α=0.005 (0.05/9)

Note. Zero-order Pearson correlations are above the diagonal, and Pearson partial correlations are below the diagonal (age and Raven matrices controlled). One-tailed p values are reported in parentheses. Log Bayes factors are reported in square brackets.
$^{*}p$ < Bonferroni-corrected $\alpha = 0.01$ (0.05/3).

$p = .09$, LBF = −0.23; Fig. 3D–F, red regression lines). For adult participants, no condition correlated with age (all *ps* > .05). These results suggest that at around 7 years of age, 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.

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 kinds 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 children's subitizing limits for sequential stimuli positively correlated between each other; children with higher subitizing limits for sequential sequences of visual events also had higher limits for sequences of tones ($r = .28$, $p = .003$, LBF = 0.54) (Table 4). However, subitizing for simultaneous stimuli did not correlate with any of the sequential conditions (dots vs. flashes: $r = .07$, $p = .23$, LBF = −1; dots vs. tones: $r = .21$, $p = .02$, LBF = −0.19). The positive significant correlation between sequential stimuli was robust given that it remained significant even when controlling for the effects of age and nonverbal IQ simultaneously ($r_p = .28$, $p = .006$) (Table 4, below diagonal). No significant correlation was found for adult participants (dots vs. flashes: $r = .28$, $p = .08$, LBF = −0.28; dots vs. tones: $r = .27$, $p = .09$, LBF = −0.92; flashes vs. tones: $r = .07$, $p = .64$, LBF = −0.89). These analyses are in line with those previously reported in this article to support the idea of two different systems for simultaneous and sequential subitizing.

Correlation between subitizing limits and mathematical abilities

We then went on to investigate whether subitizing limits correlated 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 correlated with any of the mathematical scores (overall indexes for mental addition and semantic tasks) in either children or adults (Table 5). Besides the high *p* values, Bayes factors clearly spoke in favor of the null hypothesis, with values less than −0.5 considered substantial evidence in favor of the null hypothesis and values less than −1 considered strong evidence. It is important to note that we recently showed the very same math scores correlated with spatial simultaneous numerosity estimation and discrimination precision levels for numerosities higher than the subitizing range (see Fig. 6 in Anobile et al., 2018). This clearly shows that the lack of correlation observed here does not depend on problems with measuring mathematical abilities.

We then investigated whether a significant correlation may have emerged by considering response speed or accuracy in the math tasks independently. We again performed 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 Weber fractions). The results are reported in Table 6. It is clear that no significant correlations between subitizing and mental addition proficiency emerged in any dimension or group of participants.

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 considered 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 differed among these three subpopulations, we ran a 3 (Clusters: 1, 2, or 3) × 3 (Subitizing Capacity: dots, flashes, or tones) repeated-measures ANOVA. The interaction between factors was found to be not significant, $F(4, 283) = 0.54$, $p = .70$, 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 combi-

Table 5
Zero-order Pearson correlations between subitizing capacities and math abilities in children and adults.

	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

Colour scale: 0,00; 0,08; 0,16; 0,24; 0,32; 0,41; 0.5

Note. Two-tailed p values are reported in parentheses. Adult Bonferroni-corrected α = .008 (.05/6); child Bonferroni-corrected α = .005 (.05/9). Log Bayes factors are reported in square brackets.

nations with a series of post hoc t tests. These analyses also confirmed that none of the subitizing measures differed across participant clusters (min p = .06, Bonferroni-corrected alpha = .005).

We then applied a more extreme approach; we split the child sample into two subsamples, collapsing higher and lower math-skilled participants. Because the performance on two symbolic math tasks correlated well between each other (r = .43, p < .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), and those below the 15th percentile were assigned to the "low-math" group (N14, average z score = −0.79). With a nonparametric sample-with-replacement bootstrap technique (10,000 iterations), we built the two math distributions shown in Fig. 5; on each iteration, math z scores were resampled (with replacement) separately for each math group, and the average z scores were computed. At the same time, we calculated subitizing limits separately for each stimulus condition and math group. Fig. 5C–E show 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 number of times that, on each of the 10,000 iterations, the difference between

Table 6
Zero-order Pearson correlations between subitizing capacities and mental addition proficiency in children and adults.

	Dots	Flashes	Tones
	Children		
Speed (s)	0.06 (0.5) [−1.8]	−0.13 (0.17) [−1.3]	−0.1 (0.29) [−1.5]
Precision (Wf)	−0.13 (0.21) [−1.3]	−0.12 (0.22) [−1.16]	−0.17 (0.08) [−0.55]
	Adults		
Speed (s)	−0.07 (0.66) [−1.5]	−0.16 (0.32) [−1.13]	−0.19 (0.23) [−0.9]
Precision (Wf)	−0.16 (0.33) [−1.1]	−0.34 (0.03) [0.5]	0.04 (0.81) [−1.5]

Note. Two-tailed p values are reported in parentheses. Log Bayes factors are reported in square brackets. Bonferroni-corrected α = .008 (.05/6). Wf, Weber fraction.

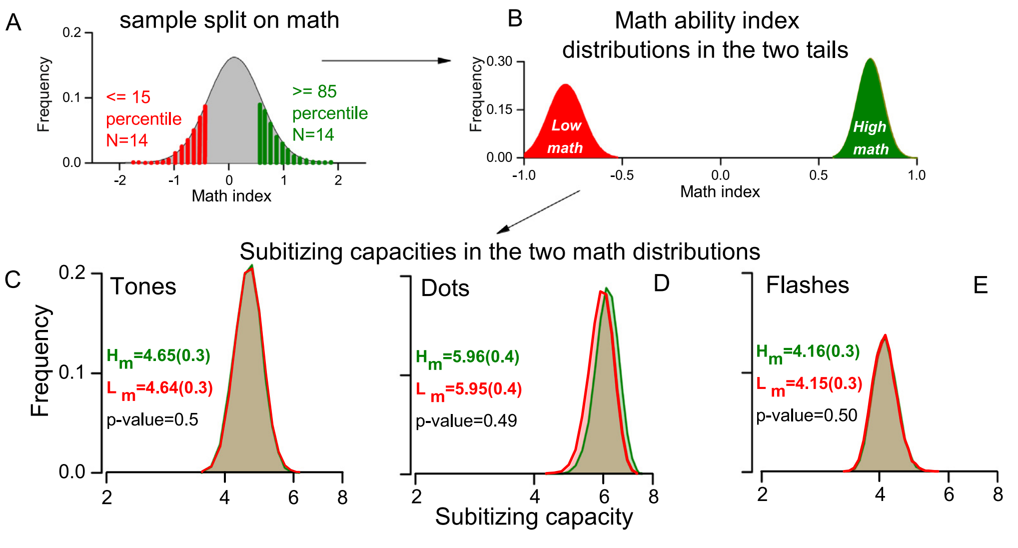

Fig. 5. Subitizing limits for "low-math" and "high-math" children. (A) Frequency distribution of child math scores, with tails (14 each) falling in the selected percentile ranges: ≥85th percentile, green, high math; ≤15th percentile, red, low math. (B) Child math scores frequency distributions inside the two percentile samples. (C–E) Subitizing limit distributions separated for the two math samples (H_m, high math, green; L_m, low math, red). Values reflect average and standard deviation (in parentheses) subitizing capacities and associated one-tail p values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the averaged subitizing capacities were higher than zero (one-tailed p value). All p values (reported in Fig. 5C–E) were near .50, 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; sample size was $n = 5$ and $n = 6$ for ≤5% and ≥95%, respectively) or the 10% tails (10–90 percentile; $n = 11$ for both ≤10% and ≥90% groups). Even in these cases, none of the subitizing measures differed between groups (all $ps \approx .50$). We then checked whether an automatized analysis might have identified the same group of participants as 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 5th percentile of math proficiency distribution (which, as shown above, showed a subitizing range rather identical to all other participants, even to those in the 95th percentile of the highest match scores).

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 nonverbal reasoning abilities (Raven matrices). We found that (a) regardless of stimuli sensory modality and presentation format (simultaneous or temporal), adults subitized one item more than children; (b) subitizing limits for simultaneous stimuli did not change between 7 and 11 years of age, whereas subitizing limits for sequential stimuli (particularly in the case of visual flashes) significantly improved during this period; (c) simultaneous and sequential subitizing limits did not correlate with each other, but sequential subitizing for visual and auditory stimuli did correlate, even over the age range of 7–11 years; and (d) in neither group of participants 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 syndrome and Down syndrome (Paterson, Girelli, Butterworth, & Karmiloff-Smith, 2006), and 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 affected 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 the Bayes factor. LBFs for correlations between subitizing limits and math skill all were clearly negative and mainly near −1 (Table 5, square brackets), indicating strong evidence in favor of the null hypothesis of zero correlations (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 the link between math and subitizing occurs only when counting is used.

Another important point to consider 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 slopes than control children, suggesting that these participants 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 Fig. 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 those of control children. 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 participants' ability to subitize.

As mentioned above, no causal studies have investigated the link between math and subitizing. Curiously, the only cognitive capacity that was causally related to subitizing was not numerical. Green and Bavelier (2003), in their seminal article, 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 et al., 2010; Burr, Anobile, & Turi, 2011; 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 current study, we reanalyzed recent data collected for other purposes. In a previous study, we only considered data in the estimation range, carefully avoiding subitizing (Anobile et al., 2018), and detected a good correlation between children's 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; therefore, the lack of correlation found here cannot be accounted for by difficulties in measurement of arithmetic abilities. In any case, because mathematics is not a single concept, the possibility still remains open that certain skills may be exclusively linked to the ANS and others may be linked to the subitizing system. In previous reports, our group has proposed that 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, the precision level in the subitizing range approaches that in the estimation range (Anobile et al., 2012; Burr et al., 2010). Moreover, although 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 working memory 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 current 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 capacities 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 an LBF near 0.5, robust enough to survive Bonferroni correction, and remained significant even when controlling for the important covariates of age and nonverbal reasoning scores (Raven matrices). 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. Because 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. Because 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 hypothesized differentiation is difficult to say, but we may speculate that it should reflect gradually reduced cross-talk between general domain skills across different sensory modalities (e.g., 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 proof 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 complementary way to test for our current results, it would be interesting to look for neuropsychological dissociations in patients with brain lesions given that 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 both simultaneous (dots) and sequential (sounds and flashes) stimuli shows the classical subitizing performance advantage, the stimuli may be subserved by separate systems. Furthermore, subitizing limits for dots and flashes, as well as for sounds, do not seem to be related to numerical abilities (at least with those measured in the current study).

Acknowledgments

This research was funded by the Italian Ministry of Health, by the Region of Tuscany under the project "Ricerca Finalizzata" (Grant No. GR-2013-02358262 to G.A.), and by the European Research Council (FP7-IDEAS-ERC) under the project "Early Sensory Cortex Plasticity and Adaptability in Human Adults—ECSPLAIN" (Grant No. 338866). G.A. was partially funded by FP7-IDEAS-ERC (Grant No. 338866).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2018.09.017.

References

Agrillo, C., Miletto Petrazzini, M. E., & Bisazza, A. (2017). Numerical abilities in fish: A methodological review. *Behavioral Processes, 141*, 161–171.
Anobile, G., Arrighi, R., Castaldi, E., Grassi, E., Pedonese, L., Moscoso, P. A. M., & Burr, D. C. (2018). Spatial but not temporal numerosity thresholds correlate with formal math skills in children. *Developmental Psychology, 54*, 458–475.
Anobile, G., Arrighi, R., Togoli, I., & Burr, D. C. (2016). A shared numerical representation for action and perception. *eLife, 5*. https://doi.org/10.7554/eLife.16161.
Anobile, G., Castaldi, E., Turi, M., Tinelli, F., & Burr, D. C. (2016). Numerosity but not texture density correlates with math ability in children. *Developmental Psychology, 52*, 1206–1216.
Anobile, G., Stievano, P., & Burr, D. C. (2013). Visual sustained attention and numerosity sensitivity correlate with math achievement in children. *Journal of Experimental Child Psychology, 116*, 380–391.
Anobile, G., Turi, M., Cicchini, G. M., & Burr, D. C. (2012). The effects of cross-sensory attentional demand on subitizing and on mapping number onto space. *Vision Research, 74*, 102–109.
Arp, S., & Fagard, J. (2005). What impairs subitizing in cerebral palsied children? *Developmental Psychobiology, 47*, 89–102.
Arp, S., Taranne, P., & Fagard, J. (2006). Global perception of small numerosities (subitizing) in cerebral-palsied children. *Journal of Clinical and Experimental Neuropsychology, 28*, 405–419.
Arrighi, R., Lunardi, R., & Burr, D. (2011). Vision and audition do not share attentional resources in sustained tasks. *Frontiers in Psychology, 2*. https://doi.org/10.3389/fpsyg.2011.00056.
Arrighi, R., Togoli, I., & Burr, D. C. (2014). A generalized sense of number. *Proceedings of the Royal Society B: Biological Sciences, 281*. https://doi.org/10.1098/rspb.2014.1791.
Ashkenazi, S., & Henik, A. (2012). Does attentional training improve numerical processing in developmental dyscalculia? *Neuropsychology, 26*, 45–56.
Ashkenazi, S., Mark-Zigdon, N., & Henik, A. (2013). Do subitizing deficits in developmental dyscalculia involve pattern recognition weakness? *Developmental Science, 16*, 35–46.
Askenazi, S., & Henik, A. (2010). Attentional networks in developmental dyscalculia. *Behavioral and Brain Functions, 6*. https://doi.org/10.1186/1744-9081-6-2.
Biancardi, A., Bachmann, C., & Nicoletti, C. (2016). *Batteria per la discalculia evolutiva (BDE2)*. Trento, Italy: Centro Studi Erickson.
Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436.
Bruandet, M., Molko, N., Cohen, L., & Dehaene, S. (2004). A cognitive characterization of dyscalculia in Turner syndrome. *Neuropsychologia, 42*, 288–298.
Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19*, 273–293.
Burr, D. C., Anobile, G., & Turi, M. (2011). Adaptation affects both high and low (subitized) numbers under conditions of high attentional load. *Seeing and Perceiving, 24*, 141–150.

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision, 10*. https://doi.org/10.1167/10.6.20.
Camos, V., & Tillmann, B. (2008). Discontinuity in the enumeration of sequentially presented auditory and visual stimuli. *Cognition, 107*, 1135–1143.
Carey, S. (2002). Cognitive foundations of arithmetic: Evolution and ontogenesis. *Mind and Language, 16*, 37–55.
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. *Research in Developmental Disabilities, 35*, 27–35.
Cipolotti, L., Butterworth, B., & Denes, G. (1991). A specific deficit for numbers in a case of dense acalculia. *Brain, 114*, 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. *Developmental Science, 13*, 120–133.
Dacke, M., & Srinivasan, M. V. (2008). Evidence for counting in insects. *Animal Cognition, 11*, 683–689.
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. *Journal of Experimental Child Psychology, 103*, 186–201.
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. *Cognitive Neuropsychology, 20*, 487–506.
Desoete, A., & Grégoire, J. (2006). Numerical competence in young children and in children with mathematics learning disabilities. *Learning and Individual Differences, 16*, 351–367.
Ditz, H. M., & Nieder, A. (2015). Neurons selective to the number of visual items in the corvid songbird endbrain. *Proceedings of the National Academy of Sciences of the United States of America, 112*, 7827–7832.
Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*, 307–314.
Ferrand, L., Riggs, K. J., & Castronovo, J. (2010). Subitizing in congenitally blind adults. *Psychonomic Bulletin & Review, 17*, 840–845.
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 Research, 1190*, 132–142.
Gray, S. A., & Reeve, R. A. (2014). Preschoolers' dot enumeration abilities are markers of their arithmetic competence. *PLoS One, 9*(4), e94428.
Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*, 534–537.
Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 10382–10385.
Kaufman, E. L., & Lord, M. W. (1949). The discrimination of visual number. *American Journal of Psychology, 62*, 498–525.
Krajcsi, A., Szabo, E., & Morocz, I. A. (2013). Subitizing is sensitive to the arrangement of objects. *Experimental Psychology, 60*, 227–234.
Lemer, C., Dehaene, S., Spelke, E., & Cohen, L. (2003). Approximate quantities and exact number words: Dissociable systems. *Neuropsychologia, 41*, 1942–1958.
Mandler, G., & Shebo, B. J. (1982). Subitizing—An analysis of its component processes. *Journal of Experimental Psychology: General, 111*, 1–22.
McLachlan, N. M., Marco, D. J. T., & Wilson, S. J. (2012). Pitch enumeration: Failure to subitize in audition. *PLoS One, 7*(4), e33661.
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.
Nieder, A. (2016). The neuronal code for number. *Nature Reviews Neuroscience, 17*, 366–382.
Olivers, C. N. L., & Watson, D. G. (2008). Subitizing requires attention. *Visual Cognition, 16*, 439–462.
Pagano, S., Lombardi, L., & Mazza, V. (2014). Brain dynamics of attention and working memory engagement in subitizing. *Brain Research, 1543*, 244–252.
Passolunghi, M. C., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology, 88*, 348–367.
Paterson, S. J., Girelli, L., Butterworth, B., & Karmiloff-Smith, A. (2006). Are numerical impairments syndrome specific? Evidence from Williams syndrome and Down's syndrome. *Journal of Child Psychology and Psychiatry, 47*, 190–204.
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*, 380–390.
Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences, 14*, 542–551.
Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., ... Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition, 116*, 33–41.
Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition, 121*, 147–153.
Plaisier, M. A., Bergmann Tiest, W. M., & Kappers, A. M. (2009). One, two, three, many—Subitizing in active touch. *Acta Psychologica, 131*, 163–170.
Plaisier, M. A., & Smeets, J. B. (2011). Haptic subitizing across the fingers. *Attention, Perception, & Psychophysics, 73*, 1579–1585.
Plaisier, M. A., Tiest, W. M. B., & Kappers, A. M. L. (2010). Grabbing subitizing with both hands: Bimanual number processing. *Experimental Brain Research, 202*, 507–512.
Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 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*, 110–122.
Railo, H., Koivisto, M., Revonsuo, A., & Hannula, M. M. (2008). The role of attention in subitizing. *Cognition, 107*, 82–104.

Reeve, R., Reynolds, F., Humberstone, J., & Butterworth, B. (2012). Stability and change in markers of core numerical competencies. *Journal of Experimental Psychology: General, 141*, 649–666.
Repp, B. H. (2007). Perceiving the numerosity of rapidly occurring auditory events in metrical and nonmetrical contexts. *Perception & Psychophysics, 69*, 529–543.
Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science, 19*, 607–614.
Riggs, K. J., Ferrand, L., Lancelin, D., Fryziel, L., Dumur, G., & Simpson, A. (2006). Subitizing in tactile perception. *Psychological Science, 17*, 271–272.
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*, 534–536.
Schleifer, P., & Landerl, K. (2011). Subitizing and counting in typical and atypical development. *Developmental Science, 14*, 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 Development, 83*, 2028–2041.
Toll, S. W., Kroesbergen, E. H., & Van Luit, J. E. (2016). Visual working memory and number sense: Testing the double deficit hypothesis in mathematics. *British Journal of Educational Psychology, 86*, 429–445.
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.
Wang, J. J., Halberda, J., & Feigenson, L. (2017). Approximate number sense correlates with math performance in gifted adolescents. *Acta Psychologica, 176*, 78–84.
Wetzels, R., & Wagenmakers, E. J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review, 19*, 1057–1064.