Developmental Psychology
2008, Vol. 44, No. 2, 544–560

0012-1649/08/$12.00 DOI: 10.1037/0012-1649.44.2.544

The Development of Automatic Numerosity Processing in Preschoolers: Evidence for Numerosity–Perceptual Interference

Laurence Rousselle and Marie-Pascale Noël
Catholic University of Louvain

Three experiments examined developmental changes in the automatic processing of numerosity and perceptual information using a nonsymbolic numerical Stroop paradigm. In Experiments 1 and 2 (E1 and E2), 4-, 5-, and 6-year-olds had to compare the numerosities or the total filled areas of collections of dots (E1) or bars (E2) varying along both dimensions. Experiment 3 replicated E2's results in 3-, 4-, and 5-year-olds. Results demonstrated the existence of reciprocal influences between numerical and perceptual information beginning at age 3. Moreover, the irrelevant perceptual influences remained stable throughout development, whereas the sensitivity to irrelevant numerical cues tended to increase with age despite children's growing inhibition capacities. No significant correlation could be found between these developmental changes and the acquisition of counting knowledge.

Keywords: automaticity, numerosity, counting, magnitude, number comparison

Supplemental materials: http://dx.doi.org/10.1037/0012-1649.44.2.544.supp

Numerosity is one of the most abstract dimensions of our environment. Unlike color, shape, or texture, numerosity is not a property of individual objects but a property of collections, which might include extremely different entities. For example, three toys, three jumps, and three sounds have nothing else in common other than their "threeness," a numerical concept that can be compared across modes and modalities of presentation. In modern societies, the use of verbal and Arabic numerals to represent numerosities further enhances the abstractness of numerical concepts, as only one symbol is used to designate a set of things.

In spite of their abstract nature, it is quite remarkable to see how easily people perceive, use, and manipulate numbers in everyday life. In adulthood, the relationship between an Arabic symbol and the meaning it conveys is so natural that adults automatically process digits' numerical meaning. Indeed, numerous studies showed that digits' magnitude modulates adults' performance in tasks that do not require any kind of number magnitude processing (Dehaene & Akhavein, 1995; den Heyer & Briand, 1986; Morin, DeRosa, & Stultz, 1967). For example, when adults have to compare two simultaneously presented digits on their physical size, they exhibit a typical *congruity effect*: They are faster at comparing digits' size when physical and numerical magnitudes are congruent (i.e., the larger digit in physical size is the larger in number: 2 vs. **3**) than when both dimensions are incongruent (i.e., the larger digit in physical size is the smaller in number: **2** vs. 3; Henik & Tzelgov, 1982; Tzelgov, Meyer, & Henik, 1992). This effect indicates that numerical information influences physical size comparison, even though this processing is not part of the task requirement. Under these conditions, number magnitude processing has been said to be automatic, as it is fast, effortless, and autonomous; that is, it starts and continues to completion without intention and monitoring (see Tzelgov & Ganor-Stern, 2005; Zbrodoff & Logan, 1986, for discussions).

In children, the first evidence for automatic semantic processing comes from an experiment requiring children from different age groups (mean age 5.8–11.0 years) and adults to judge whether two digits were the same or not (Duncan & McFarland, 1980). Although this task only necessitated a basic perceptual judgment, reaction times in all age groups decreased as the numerical distance between digits increased, suggesting that children as young as 6 years old automatically accessed number meaning. However, results later obtained by Girelli, Lucangeli, and Butterworth (2000) called this conclusion into question. In their study, first, third, and fifth graders and adults (mean age: 6.6, 8.4, 10.3, and 23 years, respectively) compared pairs of digits on their physical size while the congruity between physical and numerical dimensions (congruent: 2 – **3** vs. incongruent: **2** – 3 vs. neutral: **2** – 2) and the numerical distance between digits (1 vs. 5 units) were manipulated. In this paradigm, performance in the physical size comparison was unaffected by the numerical congruity before the third grade and by the numerical distance before the fifth grade, sug-

Laurence Rousselle and Marie-Pascale Noël, Catholic University of Louvain, Psychology Department, Cognitive Development Laboratory, Louvain-la-Neuve, Belgium.

The authors are supported by the National Fund for Scientific Research of Belgium (FNRS). The research has also been supported by two Concerted Research Actions of the Belgian French Community (Conventions 01/06-267 and 05/10-327), an FRSM fund (Fund for Scientific and Medical Research, Convention 3.4607.04 F), and a Marie Curie Research Training Network from the European Economic Community (MRTN-CT-2003-504927, Numbra project). We thank Fanny Tsolakidis and Sophie Debroux for contribution to the participants' screening and to the data collection procedure. We are grateful to the children, parents, headmasters, and teachers of the participating schools.

Correspondence concerning this article should be addressed to Marie-Pascale Noël, Cognitive Development Laboratory, Psychology Department, Catholic University of Louvain, 10 place C. Mercier, 1348, Louvain-la-Neuve, Belgium. E-mail: marie-pascale.noel@uclouvain.be

gesting that automatic number processing starts to emerge only later in development. Nevertheless, by testing first graders at different moments of the school year, Rubinsten, Henik, Berger, and Shahar-Shalev (2002) found the presence of a congruity effect in children from the end of first grade, providing evidence for automatic number processing beginning at age 7.

Altogether, these results suggest that the automaticity of number magnitude processing develops gradually over the course of learning by the time children enter school. But, would it not be possible to detect earlier signs of automatic number magnitude processing using nonsymbolic materials such as collections of elements instead of Arabic symbols? Many studies in the literature have indeed shown that younger children, infants, and even neonates are able to discriminate the numerosity of collections or sequences of visual (e.g., Antell & Keating, 1983; Canfield & Smith, 1996; Starkey, Spelke, & Gelman, 1990; Strauss & Curtis, 1981; Wood & Spelke, 2005b; Wynn, 1996) and auditory stimuli (Bijeljac-Babic, Bertoncini, & Melher, 1993; Lipton & Spelke, 2003, 2004). Most remarkably, infants are not only sensitive to the numerosity of small sets but are also able to discriminate large numerosities as long as they differ by a ratio of 1:2, that is, a ratio salient enough for a difference to be detected (Brannon, Abbott, & Lutz, 2004; Lipton & Spelke, 2003, 2004; Wood & Spelke, 2005a, 2005b; Xu, 2003; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005). This converging evidence led many authors to assume that infants are born with an innate numerosity representation system whose precision increases over development (Butterworth, 1999; Dehaene & Changeux, 1993; Gallistel & Gelman, 1992; Wynn, 1992, 1995). In this context, it should be possible to find indications of automatic numerosity processing in younger children, well before they receive formal instruction at school.

To our knowledge, only one study has provided evidence for automatic processing of number magnitude in younger children using nonsymbolic stimuli (Droit-Volet, Clément, & Fayol, 2003). In this study, 5- and 8-year-old children and adults were presented with sequences of dots and were asked to process either the sequence duration (i.e., temporal task: discriminating between long and short sequences) or the number of stimuli in the series (i.e., numerical task: discriminating between sequences containing many or few stimuli). Results revealed an asymmetrical influence of numerical over temporal discrimination: In all age groups, number interfered with temporal processing, whereas duration did not interfere with the numerical discrimination (see Dormal, Seron, & Pesenti, 2006, for similar results in adults).

The Present Research

In the present study, the development of automatic and intentional numerosity processing was examined in preschoolers from different age groups using an original nonsymbolic numerical Stroop paradigm. As the automatic activation of the irrelevant dimension depends on its relationship with the relevant dimension (Tzelgov & Ganor-Stern, 2005), the numerical comparison task was contrasted with a task requiring the processing of a closely related but non-numerical concurrent dimension. Participants were presented with two collections of filled elements and had to perform two kinds of task: a numerical comparison task, in which they had to select the larger collection in numerosity (discrete quantification), and a perceptual comparison task, in which they had to choose the collection with the larger filled area (continuous quantification). So, the relevant dimension in the first task was irrelevant in the second, and vice versa. In accordance with the Stroop paradigm principles, the congruity between the numerical and perceptual dimensions was manipulated: In congruent trials, the larger collection in numerosity also presented the larger total filled area; in incongruent trials, the larger collection in numerosity presented the smaller filled area; and finally, in neutral trials, the two collections had either the same numerosity but different area (perceptual task) or different numerosities but the same area (numerical task). The numerical and perceptual ratios between collections to be compared were also manipulated. In each task, the occurrence of an interference effect (i.e., worse performance for incongruent than for neutral trials), a facilitation effect (i.e., better performance for congruent than for neutral trials), and/or an irrelevant ratio effect was taken as a marker for the automatic processing of the irrelevant dimension.

The choice of the concurrent perceptual comparison task was motivated by recent data demonstrating the importance of perceptual cues in early quantification behavior. Indeed, several studies failed to find evidence that infants or even preschoolers discriminate between numerosities when continuous perceptual variables, and particularly, the sum of individual areas and/or circumferences, were properly controlled for (Clearfield & Mix, 1999, 2001; Feigenson, Carey, & Spelke, 2002; Rousselle, Palmers, & Noël, 2004). This led some authors to hypothesize that early quantification processes are originally undifferentiated: Infants would initially represent discrete (i.e., numerosities) and continuous (perceptual cues) quantities in terms of overall amount (Mix, Huttenlocher, & Levine, 2002; Rousselle et al., 2004) and would be able to dissociate both kinds of representation only later in the development. Without entering into this controversy, the very close relationship between numerical and perceptual properties in early quantification present them as ideal competing dimensions to be contrasted in a nonsymbolic numerical Stroop paradigm.

The main purpose of this study was to determine whether preschoolers' performance in continuous and discrete quantification tasks is respectively affected by irrelevant numerical and perceptual properties of collections. If it is, when do these influences start to occur in each task and how are they modulated throughout development? The answer to these questions may depend on several factors. First of all, developmental changes in performance should be determined by the speed and efficiency of the conflicting processing. Being a perceptual feature, surface area has been shown to be a very salient dimension for infants and 3-year-old children, as they tend to focus on this perceptual feature to the detriment of the numerical dimension when processing collections (Clearfield & Mix, 1999, 2001; Feigenson, Carey, & Hauser, 2002; Feigenson, Carey, & Spelke, 2002; Rousselle et al., 2004). This is in keeping with some observations carried out in numerical matching tasks in which young children failed to recognize the numerical equivalence between two lines of tokens of different length (Piaget & Szeminska, 1964) or between sets of dissimilar objects (Mix, 1999b), suggesting that perceptual processing predominates over numerosity processing.

Among the possible sources of developmental changes in automatic numerical and perceptual processing, one of the most significant might also be children's abilities to inhibit irrelevant influences and to focus their attention on the relevant property for the task at hand. In this respect, the occurrence of irrelevant processing in each task would be partly determined by children's ability to resist interference, a capacity known to increase during childhood along with the maturity of the frontal lobes (see Harnishfeger, 1995, for a review). As a result, the ability to process numerical and perceptual dimensions independently might be expected to improve with age along with the development of inhibition capacities.

Finally, it has been recurrently shown that the gradual acquisition of counting skills, and of cardinal concepts in particular, is positively related to the development of numerosity processing (Brannon & Van de Walle, 2001; Mix, 1999a, 1999b; Mix, Huttenlocher, & Levine, 1996; Mix, Levine, & Huttenlocher, 1997; Rousselle et al., 2004). A possible interpretation is that the mapping between the sequence of number words and their cardinal meaning promotes the abstraction of numerical relations. Beyond the cardinality principle, the abstraction and one-to-one correspondence principles of counting could also help children disregard perceptual correlated dimensions and to focus on numerosity only. These principles indeed require considering the collection as a set of distinct entities, each of which must be tagged with one number word whatever its size, shape, or other perceptual properties. The acquisition of these three ruling principles could thus promote the abstraction of number concepts and therefore favor numerosity processing development under both intentional and unintentional conditions.

Taking this as a starting point, the interaction between numerical and perceptual processing in each task could be modulated as follows. In the numerical task, the influence of perceptual on numerosity processing (as manifested by the facilitation and interference effects) is expected to be larger in younger children because of their greater sensitivity to the perceptual dimensions. With age, the acquisition of counting knowledge, supporting numerosity processing development, and the growing maturity of inhibition capacities should both contribute to reduce irrelevant perceptual influences.

By contrast, in the perceptual task, at least two possibilities could be considered. First, the automaticity of numerosity processing could be already fully developed in preschoolers. Indeed, evidence for numerosity discrimination in infants comes from nonverbal implicit paradigms (usually the habituation paradigm) in which infants were not forced or even impelled to process numerosities. Although it does not mean that they automatically processed numerosities, it nevertheless demonstrates that numerical properties are processed naturally and spontaneously very early on in the course of development. Under this hypothesis, the influence of numerical over perceptual properties should be the same across age groups (or even diminish with the growing efficiency of inhibition control). However, given the salience of the perceptual dimensions, a second possibility could be that the influence of numerical over perceptual processing emerges only gradually. In this case, the influence of numerosity should increase with age and with gradual counting mastery, despite the development of inhibition capacities.

In the present study, developmental changes in automatic and intentional numerosity processing of discrete, nonsymbolic quantities were examined in three experiments. In Experiment 1 (E1), 4-, 5-, and 6-year-olds compared the numerical and perceptual properties of collections of dots. In Experiment 2 (E2), children of the same age performed the same tasks while presented with collections of bars, a type of stimulus that enables a better control of perceptual properties. Finally, Experiment 3 (E3) reproduced E2 including a younger age group (3-year-olds) and tested the replicability of results in 4- and 5-year-olds. In all three experiments, developmental changes in intentional and automatic numerosity processing were examined in relation to the gradual acquisition of counting skills. The ability to resist interference was only assessed in E1 to confirm the growing efficiency of inhibition capacities.

Experiment 1

Method

Participants

Forty-eight Caucasian preschoolers participated in this experiment. They were drawn from a middle-class preschool in the area of Namur, Belgium. Children were evenly divided into three age groups: 4-year-olds ($n = 16$; mean age = 4 years, 1 months; SD = 2.7 months; range: 3 years, 9 months to 4 years, 5 months; 6 girls, 10 boys), 5-year-olds ($n = 16$; mean age = 5 years, 0 months; SD = 2.9 months; range: 4 years, 5 months to 5 years, 4 months; 6 girls, 10 boys), and 6-year-olds ($n = 16$; mean age = 5 years, 10 months; SD = 3.3 months; range: 5 years, 5 months to 6 years, 3 months; 7 girls, 9 boys). Four-, 5- and 6-year-olds, respectively, attended lower, intermediate, and higher level preschool classes. Generally, participants found the tasks enjoyable and asked to come back another time. The informed consent of parents and schoolteachers was obtained. No other information about participants was available.

Materials and Task Procedure

Collection comparison Stroop task. Stimuli consisted of 48 pairs of white squares (side = 8 cm) containing a variable number of black filled dots. Pairs of squares were presented on a PC screen on a dark gray background and were separated by a red fixation cross. A cat and a dog drawing were respectively placed above each array. Children were asked to show who (the cat or dog) had more dots (i.e., numerical task) or had the "more black" picture (i.e., perceptual task). Four training trials were administered before each task. Responses and latencies were recorded by the experimenter with the Cedrus RB-400 response pad (Cedrus; San Pedro, CA).

Pairs of collections varied along two dimensions: the numerosity and the total filled area, which were manipulated in three congruity conditions. In congruent trials, the larger array in number also had the larger total filled area, whereas in incongruent trials, the larger collection in number had the smaller total filled area (see Figures 1a and 1b, respectively). Neutral trials varied on one dimension at a time, depending on the task. In the numerical task, the number of dots varied but the total filled area was equated in each neutral pair (see Figure 1c). Conversely, in the perceptual task, the total filled area varied but the number of dots was the same in each neutral pair (see Figure 1d). Thus, congruent and incongruent trials were the same for both tasks, whereas the set of neutral stimuli was different for each task.

Congruent and incongruent stimuli also varied along two numerical ratios (N-ratio: 1:2 vs. 2:3) and two area ratios (A-ratio:

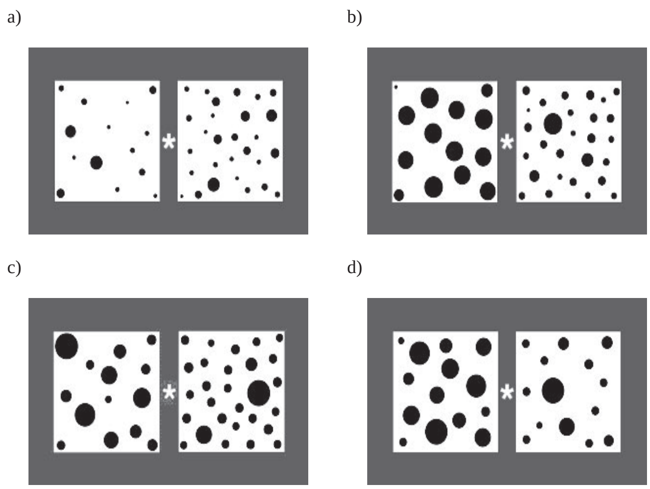

Figure 1. Examples of stimuli presented in each condition in Experiment 1: (a) congruent pair, 14 vs. 28 dots, area ratio 1:2; (b) incongruent pair, 14 vs. 28 dots, area ratio 1:2; (c) neutral pair in the numerical task, 14 vs. 28 dots, 1,500 mm^2 on both sides; (d) neutral pair in the perceptual task, 14 dots on both sides, area ratio 1:2.

1:2 vs. 2:3). As neutral pairs only varied on one dimension at a time, they varied along the N-ratios in the numerical task and along the A-ratios in the perceptual task. Given that small numerosities are supposed to be apprehended by different quantification processes from large ones (Simon, 1997; Trick & Pylyshyn, 1994), only numerosities above the subitizing range (i.e., higher than 4) were presented (from 6 to 30). Table 1 presents numerical and total filled area properties of congruent and incongruent stimuli, and Table 2 reports neutral trials properties in each task.

The total filled area was manipulated in each pair by varying the area of the individual elements. As the present stimuli were black and white, total filled area variations were completely confounded with brightness variations. To avoid the larger collection in number systematically being the one with the smaller elements, the area of the smaller and larger dots was the same in both arrays to be compared. Finally, the external perimeter of collections (formed by the most external dots) was equated for all trials.

To keep children focused on the task, only four collection pairs were presented in each experimental condition, for a total of 40 pairs presented in each task (4 pairs $\times$ 2 N-ratios $\times$ 2 A-ratios $=$ 16 congruent and 16 incongruent stimuli $+$ 4 pairs $\times$ 2 N- or A-ratios $=$ 8 neutral stimuli in each task). In each experimental condition, the larger numerosity and/or total area appeared on the right side for two stimuli and on the left side for the two others. Stimuli were presented in a fixed pseudorandom order with the constraints that no more than four consecutive trials required the same response and that no more than two stimuli from the same congruity condition appeared in succession. Each trial started with the presentation of a pair until response. The appearance of the next trial was triggered by the experimenter when the child was attentive.

Day–night task. The development of cognitive inhibition was assessed using a computerized version of the day–night task (Gerstadt, Hong, & Diamond, 1994), a task significantly related to other cognitive inhibition measures (Carlson & Moses, 2001). In the naming condition, children had to say "day" when shown the picture of a yellow sun in a blue sky and "night" when shown the picture of a white moon in a dark sky. This condition was then followed by an interfering condition in which they had to say "day" when shown the moon picture and "night" when shown the sun picture. This condition required children to inhibit their prepotent tendency to name the picture. Each condition started with 4 training trials, followed by 16 test trials. Stimuli were presented in a fixed pseudorandom order with the constraint that the same picture never appeared in three consecutive trials. Responses and

Table 1
Numerosity Pairings and Total Filled Area Properties of Congruent and Incongruent Stimuli

Congruent stimuli		Incongruent stimuli	
N-ratio 1:2	N-ratio 2:3	N-ratio 1:2	N-ratio 2:3
A-ratio 1:2			
6–12 (250–500 mm^2)	9–6 (500–250 mm^2)	6–12 (500–250 mm^2)	9–6 (250–500 mm^2)
16–8 (2,000–1,000 mm^2)	10–15 (1,000–2,000 mm^2)	16–8 (375–750 mm^2)	10–15 (750–375 mm^2)
14–28 (375–750 mm^2)	24–16 (750–375 mm^2)	14–28 (2,000–1,000 mm^2)	24–16 (1,000–2,000 mm^2)
30–15 (3,000–1,500 mm^2)	20–30 (1,500–3,000 mm^2)	30–15 (1,500–3,000 mm^2)	20–30 (3,000–1,500 mm^2)
A–ratio 2:3			
6–12 (250–375 mm^2)	9–6 (375–250 mm^2)	6–12 (375–250 mm^2)	9–6 (250–375 mm^2)
16–8 (1,500–1,000 mm^2)	10–15 (1,000–1,500 mm^2)	16–8 (1,000–1,500 mm^2)	10–15 (1,500–1,000 mm^2)
14–28 (1,500–2,250 mm^2)	24–16 (2,250–1,500 mm^2)	14–28 (2,250–1,500 mm^2)	24–16 (1,500–2,250 mm^2)
30–15 (3,000–2,000 mm^2)	20–30 (2,000–3,000 mm^2)	30–15 (2,000–3,000 mm^2)	20–30 (3,000–2,000 mm^2)

Note. Total filled areas reported in parentheses are given as an example. A-ratio $=$ area ratio; N-ratio $=$ numerical ratio.

Table 2
Numerosity Pairings and Total Filled Area Properties of Neutral Stimuli in Each Task

Perceptual task		Numerical task	
A-ratio 1:2	A-ratio 2:3	N-ratio 1:2	N-ratio 2:3
6–6	6–6	12–6	9–6
(250–500 mm^2)	(250–375 mm^2)	(250–250 mm^2)	(250–250 mm^2)
14–14	14–14	8–16	10–15
(2,000–1,000 mm^2)	(1,500–1,000 mm^2)	(1,000–1,000 mm^2)	(1,000–1,000 mm^2)
15–15	15–15	28–14	24–16
(375–750 mm^2)	(1,500–2,250 mm^2)	(1,500–1,500 mm^2)	(1,500–1,500 mm^2)
30–30	30–30	15–30	20–30
(3,000–1,500 mm^2)	(3,000–2,000 mm^2)	(3,000–3,000 mm^2)	(3,000–3,000 mm^2)

Note. Total filled areas reported in parentheses are given as an example. A-ratio = area ratio; N-ratio = numerical ratio.

latencies were recorded by the experimenter on the Cedrus RB-400 response pad. Each correct response was credited with 1 point.

For each participant, an interference effect was calculated as the difference between the percentage of correct response in the naming and interfering conditions (Naming – Interference). An interference index was also calculated on the basis of correct median reaction times using the following formula: (Naming – Interference) / Naming. Lower inhibition scores and indexes indicate greater inhibitory control.

Counting tasks. Two tasks were administered to assess counting development. The trade task was designed to assess the nonverbal understanding of the abstraction and the one-to-one correspondence principles. In this task, children were given tokens that they were invited to trade for farm animals: "You have coins and I have farm animals to sell. If I give you one animal, you have to give me one coin." Two practice trials ensued in which the child was given a big and little animal to make sure he understood that the size does not matter. The task was then administered in three steps starting with numerosities from 1 to 4, followed by numerosities from 5 to 8 and then from 9 to 12. This task could be performed by counting or simply by giving one token for one animal. In either case, correct answers required making abstraction of the nature of the animals and applying a strict one-to-one correspondence. Trials were administered in a nonconsecutive order within each step (e.g., 2, 1, 4, and 3) until the child made errors on three successive numerosities. The trade score corresponded to the highest number of animals correctly traded.

Modeled on a task developed by Wynn (1992), the point-to-*x* task was used to test whether the child knew the precise numerosity a number word refers to, namely, its cardinal meaning. Children were shown three cards, each presenting four arrays with different number of frogs or bananas. The first card represented numerosities from 1 to 4, the second, numerosities from 5 to 8, and the third, numerosities from 9 to 12. Children were asked to point to the picture depicting a number (*n*) of items. For each card, numerosities were requested in a nonconsecutive order (e.g., 2, 1, 4, and 3). Trials were administered until the child made errors on three consecutive numerosities. Each correct response was credited with 1 point.

Testing procedure

Children were tested individually in a quiet room within their school. Testing took place in May, at the end of the school year, and was completed in two 20-min sessions. One session started with the numerical collection comparison Stroop task and ended with the day–night task. The other started with the perceptual collection comparison Stroop task, followed by the trade task and then the point-to-*x* task. The order of the sessions was counterbalanced across participants. Regular breaks were made during the session to maintain optimal level of attention.

Results

Collection Comparison Stroop Tasks

Table 3 reports mean percentages of correct responses by task and congruity condition for each age group. These percentages were very low in some cells of the table, especially for incongruent trials (from 62.89 to 64.84% in the numerical task and from 33.59 to 50% in the perceptual task). Given the small number of items in each experimental condition and the large number of errors, analyses were carried out on accuracy data only. However, accuracy data did not fit the sphericity assumption of the repeated-measures analysis of variance (ANOVA) statistical model (Mauchly's test). As multivariate analyses of variance (MANOVA) do not require the sphericity of the variance–covariance matrix (Howell, 2001, p. 519), percentages of correct responses in each task were analyzed in a 3×3 repeated-measures MANOVA (Wilks' lambda test statistic) with condition (congruent, neutral, incongruent) as a within-subjects factor and age group (4-year-olds, 5-year-olds, 6-year-olds) as a between-subjects factor. Multiple pairwise comparisons were adjusted using the Bonferroni correction. An alpha level of .05 was used for all statistical tests.[1]

Figure 2 represents percentages of correct responses by task and congruity condition for each group (see also Table 3). In the numerical task, the main effects of condition and age group were significant, $F(2, 44) = 32.95$, $p < .001$, $\eta^2 = .60$, and $F(2, 45) = 3.94$, $p < .05$, $\eta^2 = .15$, respectively. However, the two factors did not interact with each other, $F(4, 88) < 1$, $p > .10$. Decomposition into contrasts revealed that all conditions differed significantly from one another (ps $< .001$) and that 4-year-olds tended to

[1] Complete full factorial analyses including A-ratio and N-ratio effects are available in supplemental materials for all three experiments. They were not presented here, as ratios had only minimal or no influence on the pattern of results.

Table 3
Experiment 1: Mean Performance by Task for Each Age Group

Task	4 years			5 years			6 years			Total		
	M	*SD*	CI (95%)	*M*	*SD*	CI (95%)	*M*	*SD*	CI (95%)	*M*	*SD*	CI (95%)
Perceptual Stroop task												
Congruent	73.1	15.9	64.9–81.2	75.4	19.4	67.2–83.6	83.6	12.7	75.4–91.8	77.3	16.5	72.6–82.1
Neutral	66.4	23.2	56.1–76.8	70.3	20.4	60.0–80.7	77.3	17.8	67.0–87.7	71.4	20.6	65.4–77.3
Incongruent	50.0	18.5	39.8–60.2	48.8	21.1	38.7–59.0	33.6	20.9	23.4–43.8	44.1	21.2	38.3–50.0
Total	63.2	11.5	57.8–68.5	64.8	10.8	59.5–70.2	64.8	9.2	59.5–70.2	64.3	10.4	61.2–67.3
Numerical Stroop task												
Congruent	79.3	14.4	73.1–85.5	91.0	10.9	84.8–97.2	90.6	11.4	84.4–96.8	87.0	13.3	83.4–90.6
Neutral	69.5	19.9	61.1–78.0	82.0	9.1	73.6–90.5	83.6	19.2	75.1–92.1	78.4	17.6	73.5–83.3
Incongruent	62.9	13.0	55.0–70.8	64.8	16.8	57.0–72.7	63.7	16.8	55.8–71.5	63.8	15.3	59.3–68.3
Total	70.6	10.2	65.5–75.7	79.3	8.2	74.2–84.4	79.3	11.8	74.2–84.4	76.4	10.8	73.4–79.3
Day–night task												
Naming: RTs (ms)	2,087	797	1,815–2,358	1,422	259	1,159–1,685	1,336	309	1,074–1,599	1,615	608	1,462–1,768
Interference: RTs (ms)	2,479	1,211	2,077–2,880	1,844	268	1,455–2,232	1,861	457	1,472–2,249	2,061	804	1,834–2,288
Naming: % CR	90.8	13.1	86.4–95.3	96.1	5.5	91.8–100.4	97.7	3.1	93.4–101.9	94.9	8.7	92.4–97.4
Interference: % CR	63.8	23.7	54.7–72.8	92.2	17.5	83.4–101.0	96.1	4.5	87.3–104.9	84.0	22.3	78.9–89.1
Counting tasks												
Trade	4.1	2.9	2.5–5.9	6.9	3.3	5.3–8.6	8.7	3.3	7.1–10.3			
Point–to–*x*	3.8	2.0	2.5–5.1	6.4	3.4	5.2–7.7	5.8	1.3	4.6–7.0			

Note. CI = confidence interval; RTs = median reaction times (based on correct responses only); CR = correct responses.

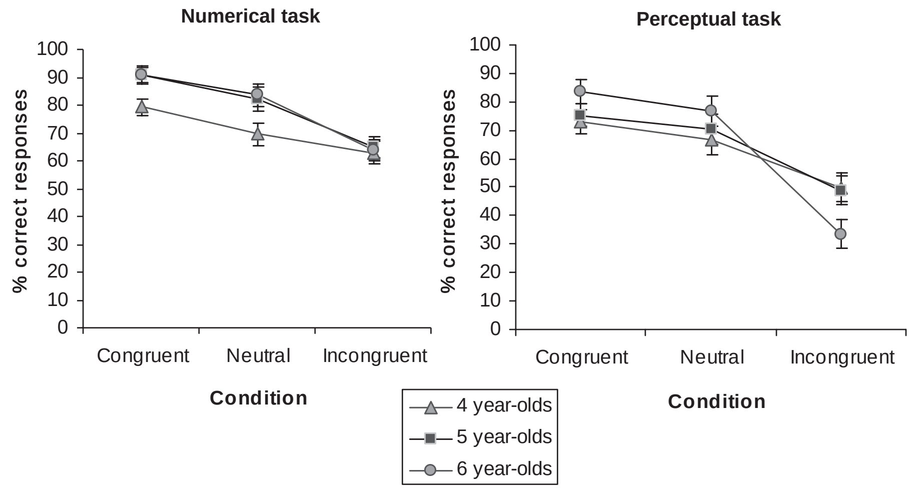

Figure 2. Experiment 1: Congruity effect in the perceptual and numerical comparison tasks; percentages of correct responses and standard errors for each age group.

perform significantly lower than 5- and 6-year-olds (ps = .057). The two older groups did not differ from each other (p = 1).

In the perceptual task, the main effect of congruity condition was significant, $F(2, 44) = 35.25$, $p < .001$, $\eta^2 = .62$, but the three age groups did not differ from each other, $F(2, 45) < 1$. The percentages of correct responses were significantly lower for incongruent trials than for congruent and neutral stimuli (i.e., congruity and interference effects, ps $< .001$, see Table 3). No reliable difference was found between congruent and neutral trials (i.e., facilitation effect, $p = .13$). The interaction between condition and age group was marginally significant, $F(4, 88) = 2.16$, $p = .08$, $\eta^2 = .09$. As shown in Figure 2, the condition effect was significant in all age groups, Fs $(2, 14) > 5.44$, ps $< .05$, but tended to be slightly greater in 6-year-olds.[2]

Two-tailed t-tests were used to compare the percentages of correct responses for each age group with the percentage expected by chance (i.e., 50%) in each task and congruity condition (see Table 3). In the numerical task, all age groups performed above chance in all congruity conditions, ts$(15) > 3.25$, ps $< .01$. In the perceptual task, all age groups performed above chance level on congruent and neutral trials, ts$(15) > 2.84$, ps $< .05$. In the incongruent condition, however, children from the two youngest age groups performed at chance, $t(15) = .00$ and $-.22$, respectively, ps $> .10$, whereas 6-year-olds performed significantly under chance level, $t(15) = -3.14$, $p < .01$, suggesting that they were highly influenced by the irrelevant numerical dimension.

Day–Night Task

Table 3 reports mean reaction times and mean percentages of correct responses in the naming and interfering conditions of the day–night task for each age group. A 2 × 3 repeated-measures ANOVA with condition (naming vs. interfering) as a within-subjects factor and age group (4 year-olds, 5 year-olds, 6 year-olds) as a between-subjects factor was run on correct median reaction times and percentage of correct responses, respectively. Both reaction times and accuracy analyses yielded main effect of age group, Fs $(2, 45) > 8.74$, ps $< .001$: Four -year-olds performed slower and less accurately than 5- and 6-year-olds (ps $<$ .005), whereas the two older groups had similar performance (ps = 1). Moreover, children performed faster and more accurately in the naming than in the interfering condition, Fs $(1, 45) > 15.70$, ps $< .001$. In the accuracy analysis, this condition effect decreased significantly with age, indicating an increase of inhibitory control with children growing, $F(2, 45) = 8.86$, $p = .001$, $\eta^2 = .28$; for reaction times analysis: $F(2, 45) < 1$.

Pearson correlations were also calculated between day–night interference effects (based on correct response), day–night interference indexes (based on relative latencies), and the interference effect in the perceptual and the numerical task. However, none of the correlations was significant whether or not age was partialled out (ps $> .10$), confirming the independence between the development of inhibition and of automatic perceptual and numerical processing.

Counting Tasks

In this experiment, the last issue was to assess the relationship between the acquisition of counting skills and the development of numerosity processing under intentional and unintentional conditions. Therefore, Pearson correlations were calculated between age, counting scores (see Table 3), percentages of correct responses for neutral stimuli in the numerical task (i.e., intentional numerical processing), and interference effects in the perceptual task (i.e., unintentional numerical processing). As shown in Table 4, all correlations with age were significant. Scores from the two counting tasks also correlated significantly with each other, even when age was partialled out. Finally, trade scores correlated significantly

[2] To assess how the effect of task order modulated the congruity effect size across age, a 2 × 3 MANOVA with order (numerical task first, perceptual task first) and age group (4-year-olds, 5-year-olds, 6-year-olds) as between-subjects factors was run in each task on the congruity effect. In both tasks, neither the effect of order, $F(1, 42) < 1.05$, $p > .05$, nor the Order × Age group interaction, $F(2, 42) < 1.58$, $p > .05$, was significant.

Table 4
Experiment 1: Correlations Between Counting Tasks, Performance for Neutral Stimuli in the Numerical Task (NT), and Interference Effect in the Perceptual Task (PT)

Variable	1	2	3	4	5
1. Age	—				
2. Trade	$.54^{***}$	—	$.34^{*}$	.24	−.06
3. Point-to-*x*	$.32^{*}$	$.44^{**}$	—	.14	−.10
4. NT: Neutral trials	$.28^{*}$	$.35^{*}$	.22	—	−.10
5. PT: Interference	$.33^{*}$	.13	.02	.01	—

Note. The area below the diagonal presents Pearson correlation coefficients and the above the diagonal reports partial correlations controlling for age.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

with performance for neutral stimuli in the numerical task, but this last correlation was no longer significant when age was partialled out. All other correlations were nonsignificant whether or not age was controlled for.

Discussion

E1 suggests that children as young as 4 years old automatically process perceptual and numerical properties of collections. In the numerical task, although the influence of irrelevant perceptual dimensions was expected to be larger in younger children and to decrease with age along with the development of inhibition control and counting knowledge, all age groups exhibited a similar congruity effect. Likewise, in the perceptual task, a congruity effect was detected even in the youngest age group and tended to increase slightly with age, despite children's growing ability to resist to interference.

The presence of congruity effects in both tasks indicates that irrelevant numerical and perceptual cues respectively affect continuous and discrete quantification in young children beginning at age 4. However, in spite of these reciprocal influences, some clues suggest that the automatic processing of numerical and perceptual dimensions follows different developmental trajectories. Indeed, although the influence of perceptual on numerical processing seemed to be stable throughout development in the numerical task, the sensitivity to the irrelevant numerical dimension in the perceptual task tended to increase with age, although this trend was only marginal. Furthermore, in the numerical task, all age groups were able, to some extent, to stay focused on the relevant numerical dimension as they performed above chance level even on incongruent trials. By contrast, in the incongruent condition of the perceptual task, children seemed to be increasingly influenced by irrelevant numerical properties, as the two youngest age groups performed at random, whereas 6-year-olds performed significantly under chance level. Taken together, these results suggest that the perceptual properties had relatively constant influence over numerosity processing throughout development, whereas, on the other hand, the interference of numerosity in perceptual processing tended to increase with age.

A rather unexpected result in E1 was the absence of a specific relationship between counting level and intentional or automatic numerosity processing. One possible reason for this nonsignificant result might be the type of stimuli used in E1. Although collections of dots are the most frequently used stimuli in infant studies, this kind of material does not allow to simultaneously control the total filled area and the sum of the individual circumferences (and diameters) of collections,[3] another salient perceptual dimension that naturally covaries with numerosities (Clearfield & Mix, 1999, 2001). This uncontrolled perceptual cue brings confusion on the processes at work in both Stroop tasks. Indeed, when the total filled area was equated in both collections to be compared as in the neutral condition of the numerical comparison task, numerosity still covaried with the summed circumference (and diameter) of the collection. In this case, this non-numerical dimension could have been used to succeed in the numerical task. Likewise, in the perceptual task, the congruity effect supposed to reflect the automatic processing of numerosities could in fact result from the influence of this uncontrolled non-numerical cue. This confound could have therefore led to overestimation of the contribution of numerosity processing to performance in both tasks and could explain the absence of any relationship with counting abilities. In E2, this problem was overcome by using collections of bars, a kind of stimuli enabling a better control of both total filled area and summed perimeter.

Experiment 2

Method

Participants

Forty-eight Caucasian preschoolers participated in this experiment. They were drawn from a local middle-class preschool in Brussels, Belgium, and were of middle to upper socioeconomic status. Children were evenly divided into three age groups: 4-year-olds ($n = 16$; mean age = 4 years, 0 months; $SD = 3.5$ months; range: 3 years, 5 months to 4 years, 4 months; 7 girls, 9 boys), 5-year-olds ($n = 16$; mean age = 4 years, 11 months; $SD = 3.3$ months; range: 4 years, 5 months to 5 years, 3 months; 10 girls, 6 boys), and 6-year-olds ($n = 16$; mean age = 5 years, 11 months; $SD = 5.2$ months; range: 5 years, 5 months to 7 years, 1 month; 11 girls, 5 boys). As in E1, 4-, 5-, and 6-year-olds, respectively, attended lower, intermediate, and higher level preschool classes. No other information about participants was available.

Materials and Task Procedure

Collection comparison Stroop task. Tasks and stimuli were the same as in E1, except as follows. Children were presented with collections of black filled vertical bars. As all bars were 2 mm wide, total filled area variations were perfectly confounded with total length (sum of the individual lengths) and brightness variations. Moreover, varying total filled area amounted to varying total perimeter (i.e., sum of individual perimeters) in almost identical proportions (no more than 3% of difference between total filled

[3] The area of a circle results from an exponential function of the radius (πr^2), and its circumference results from a linear function of the radius ($2r\pi$) or of the diameter ($D\pi$). Accordingly, two sets with different numbers of circles cannot present simultaneously the same total surface area and the same summed circumference or diameter.

area and total perimeter ratios). Examples of stimuli are shown in Figure 3.

To make sure that there was no confusion between numerical and perceptual instructions, a familiarization phase was conducted before each task to clearly dissociate numerical and perceptual dimensions and to draw children's attention to the relevant property for the task.[4] Familiarization trials (maximum = 8) were administered in a fixed random order until the child reached a familiarization criterion of four consecutive correct responses. All children included in the present sample met this criterion.

Counting tasks. The counting assessment was the same as in E1 but was completed with an enumeration task designed to directly assess general counting mastery. In this task, children were invited to count a set of 2 to 12 aligned frogs or bananas. Children were first presented with numerosities from 1 to 4 (Stage 1), then with numerosities from 5 to 8 (Stage 2), and finally with numerosities from 9 to 12 (Stage 3). Within each stage, numerosities were presented in a nonconsecutive order (e.g., 2, 1, 4, and 3). Trials were administered until the child made errors on three consecutive numerosities. Each correct enumeration was credited with 1 point.

Testing Procedure

The testing procedure was the same as in E1 except that the enumerating task was administered between the trade and the point-to-*x* tasks.

Results

Collection Comparison Stroop Tasks

As displayed in Table 5, mean percentages of correct responses were very low in the incongruent condition (from 39.06 to 56.64% in the numerical task and from 34.38 to 42.19% in the perceptual task). Analyses were thus carried out on accuracy data only (Wilks' lambda) as the hypothesis of sphericity could not be assumed (Mauchly's test). Percentages of correct responses in each task were analyzed in a 3 (condition: congruent, neutral, incongruent) $\times$ 3 (age group: 4-year-olds, 5-year-olds, 6-year-olds) repeated-measures MANOVA.

Figure 4 represents percentages of correct responses by task and congruity condition for each group (see also Table 5). In the numerical task, the main effects of condition and age group were significant, $F(2, 44) = 39.77$, $p < .001$, $\eta^2 = .64$, and $F(2, 45) = 6.83$, $p < .005$, $\eta^2 = .23$, respectively. All conditions differed significantly from one another (ps $< .001$), and 4-year-olds performed lower than the older groups, although this trend was only marginal in comparison to 5-year-olds ($p = .08$ and $p < .005$ for 5- and 6-year-olds, respectively). The older groups had similar performance ($p > .10$). There was no interaction between condition and age group, $F(4, 88) = 1.36$, $p > .10$.

In the perceptual task, the main effects and the interaction were significant: age group, $F(2, 45) = 7.27$, $p < .005$, $\eta^2 = .24$; condition, $F(2, 44) = 39.77$, $p < .001$, $\eta^2 = .64$; interaction, $F(4, 88) = 2.94$, $p < .05$, $\eta^2 = .12$. Four-year-olds had significantly lower performance than 5- and 6-year-olds (ps $< .05$), who did not differ from each other ($p = 1$). Moreover, there were significant congruity and interference effects (ps $< .001$) but there was no facilitation effect ($p > .10$). This condition effect was significant in all age groups and increased with age: mean congruity effect = 23%, 41%, and 56% and Fs $(2, 14) = 9.00$, 19.66, and 24.04 in 4-, 5-, and 6-year-olds, all ps $< .001$.[5]

As in E1, two-tailed *t*-tests were used to compare the percentages of correct responses for each age group with the percentage expected by chance (i.e., 50%) in each task and congruity condition. In the numerical task, all age groups performed above chance level on congruent and neutral trials, ts$(15) > 2.54$, ps $< .05$, except 5-year-olds, who performed only marginally above chance level on neutral pairs, $t(15) = 1.82$, $p = .09$. On incongruent trials, 4-year-olds performed significantly under chance level, $t(15) = -2.87$, $p < .05$, whereas the two older age groups performed at

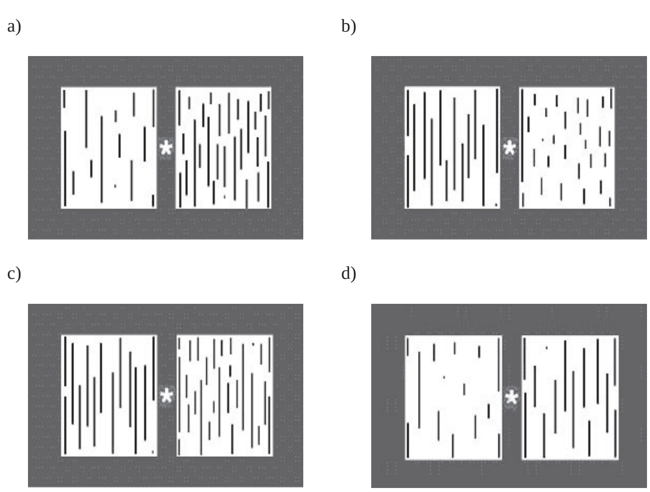

Figure 3. Examples of stimuli presented in each condition in Experiment 2: (a) congruent pair, 14 vs. 28 bars, area ratio 1:2; (b) incongruent pair, 14 vs. 28 bars, area ratio 1:2; (c) neutral pair in the numerical task, 14 vs. 28 bars, 2,000 mm^2 on both sides; (d) neutral pair in the perceptual task, 14 bars on both sides, area ratio 1:2.

[4] Participants were presented with two drawings containing a variable number of bars. In the numerical familiarization phase, they were told: "One bar can be small [showing one small bar] or tall [showing one tall bar], but it is always one bar. For us, the size doesn't matter. So, these two pictures are someway the same: there is one and one. By contrast, those two pictures are not the same: there are two bars here [showing two small bars] while there is one bar there [showing one tall bar of equivalent area as the two small bars together]. Two is more than one. And here, could you tell me where there are more bars?" In the perceptual familiarization phase, children were told: "On these pictures, there are small and large bars like pieces of chocolate. Here for example, there is only a little bit of chocolate [picture with two small bars] while there, there is a lot of chocolate [picture with one tall and one small bar]. So, if you wanted to eat a lot of chocolate, what would you choose?" For each familiarization phase, half of the trials were neutral and half were incongruent.

[5] The effect of task was assessed in each task in a 2 $\times$ 3 MANOVA with order (numerical task first, perceptual task first) and age group (4-year-olds, 5-year-olds, 6-year-olds) as between-subjects factors and congruity effect as the dependent variable. In both tasks, the effect of order was not significant, $F(1, 42) < 2.38$, $p > .05$. The Order $\times$ Age group interaction reached significance in the numerical task only, $F(2, 42) = 3.25$, $p = .05$; perceptual task, $F(2, 42) < 1$. In this task, 4- and 6-year-olds showed a larger congruity effect when starting with the perceptual task, whereas 5-year-olds showed a larger congruity effect when starting with the numerical task.

Table 5
Experiment 2: Mean Performance in the Stroop Collection Comparison Task for Each Age Group

	4 years			5 years			6 years			Total		
Task	*M*	*SD*	CI (95%)	*M*	*SD*	CI (95%)	*M*	*SD*	CI (95%)	*M*	*SD*	CI (95%)
Perceptual Stroop task												
Congruent	62.9	16.1	54.8–71.0	83.2	17.6	75.1–91.3	90.6	14.3	82.5–98.7	78.9	19.7	74.2–83.6
Neutral	66.4	16.9	56.3–76.5	77.3	23.4	67.3–87.4	83.6	19.2	73.5–93.7	75.8	20.9	70.0–81.6
Incongruent	39.8	15.8	27.7–52.0	42.2	21.2	30.1–54.3	34.4	32.3	22.2–46.5	38.8	23.8	31.8–45.8
Total	56.4	8.7	51.1–61.7	67.6	13.7	62.3–72.9	69.5	8.3	64.2–74.8	64.5	11.9	61.4–67.6
Numerical Stroop task												
Congruent	68.0	18.0	60.0–76.0	82.0	18.4	74.0–90.0	92.6	10.0	84.6–100.6	80.9	18.6	76.2–85.5
Neutral	59.4	14.8	48.0–70.8	64.1	30.9	52.7–75.4	77.3	18.9	66.0–88.7	66.9	23.4	60.4–73.5
Incongruent	39.1	15.2	28.2–50.0	56.6	21.5	45.7–67.5	55.1	26.7	44.2–66.0	50.3	22.7	44.0–56.6
Total	55.5	5.3	47.9–63.1	67.6	19.9	60.0–75.2	75.0	16.1	67.4–82.6	66.0	16.9	61.6–70.4
Counting tasks												
Enumeration	5.4	3.5	4.0–6.9	9.3	3.4	7.8–10.7	11.6	1.2	10.1–13.0			
Trade	4.7	2.9	3.1–6.3	6.8	2.8	5.2–8.3	8.5	3.7	6.9–10.1			
Point–to–*x*	3.2	1.6	2.0–4.4	5.3	2.6	4.2–6.5	7.4	2.6	6.3–8.6			

Note. CI = confidence interval.

random, $t(15) = .76$ and 1.24, respectively, *p*s > .10, suggesting that younger children were more influenced by irrelevant perceptual dimensions than older children. In the perceptual task, children from all age groups performed above chance level on congruent and neutral trials, *t*s $(15) > 3.21$, *p*s < .01. In the incongruent condition, however, the youngest and the oldest age group tended to perform significantly under chance level, $t(15) = -2.57$, $p = .02$, and $t(15) = -1.94$, $p = .07$, respectively, whereas 5-year-olds' performance did not differ significantly from chance level, $t(15) = -1.47$, $p > .10$. Although disparate, these results highlighted the strong influence of the irrelevant numerical dimension on perceptual comparisons even in the youngest age group.

Finally, a 2 (task: perceptual vs. numerical) × 2 (experiment: 1 vs. 2) × 3 (age group: 4-year-olds, 5-year-olds, 6-year-olds) ANOVA with task as the only within-subjects factor was run to compare performance across tasks when comparing collections of dots (E1) or collections of bars (E2, see Table 3 and 5). As in previous analyses, the effect of age group was significant, $F(2, 90) = 10.84$, $p < .001$, $\eta^2 = .19$. Four-year-olds scored lower than the two older groups ($p <$ .005), who did not differ from each other ($p = 1$). Moreover, the performance was significantly higher for the numerical than for the perceptual task, $F(1, 90) = 25.69$, $p < .001$, $\eta^2 = .22$, and in E1 than in E2, $F(1, 90) = 6.53$, $p < .05$, $\eta^2 = .07$. However, the significant Task × Experiment interaction, $F(1, 90) = 15.51$, $p < .001$, $\eta^2 = .15$, indicated that the advantage of dots over bars processing was only significant for the numerical task, $F(1, 94) = 12.91$, $p < .001$, $\eta^2 =$ .12; perceptual task, $F(1, 94) < 1$. Taking another point of view, this interaction also indicated that the effect of task was only significant in E1, when comparing collections of dots, $F(1, 47) = 47.66$, $p < .001$, $\eta^2 = .50$; in E2, $F(1, 47) < 1$. Other interactions were not significant, *F*s $(2, 90) < 2.64$, *p*s > .05.

Counting Tasks

As in E1, Pearson correlations were calculated to assess the relationship between age, performance in counting tasks (see Table

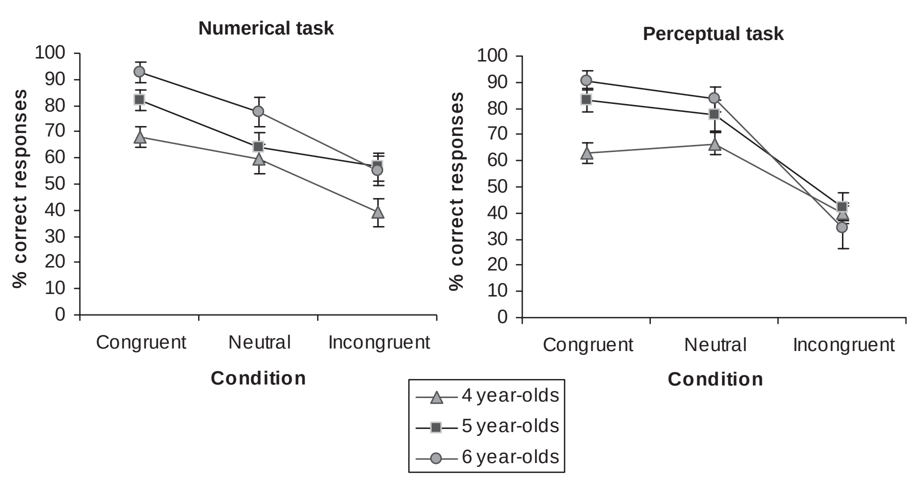

Figure 4. Experiment 2: Congruity effect in the numerical and perceptual comparison tasks; percentages of correct responses and standard errors for each age group.

Table 6
Experiment 2: Correlations Between Counting Tasks, Performance for Neutral Stimuli in the Numerical Task, and Interference Effect in the Perceptual Task

Variable	1	2	3	4	5	6
1. Age	—					
2. Enumeration	.60***	—	.22	.28†	.10	−.17
3. Trade	.47***	.44**	—	.54***	.23	.07
4. Point-to-*x*	.65***	.56***	.67***	—	.23	.16
5. N task: Neutral trials	.39**	.30*	.37*	.41**	—	−.02
6. P task: Interference	.32**	.06	.21	.33*	.11	—

Note. The area below the diagonal presents Pearson correlation coefficients, and the area above the diagonal reports partial correlations controlling for age. N task = numerical task; P task = perceptual task.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

5), percentages of correct responses for neutral stimuli in the numerical task (i.e., intentional numerical processing), and interference effects in the perceptual task (i.e., unintentional numerical processing). As displayed in Table 6, correlations with age were all significant, as were correlations between the three counting tasks, even when age was controlled for. Counting scores correlated significantly with performance on neutral trials in the numerical task, and the point-to-*x* scores correlated significantly with the interference effects in the perceptual task. However, these correlations were not significant when age was partialled out.

Discussion

Using collections of bars, E2 replicated the results obtained in E1 with collections of dots. In all age groups, performance in the numerical and perceptual tasks was affected by the congruity between the relevant and irrelevant dimensions. These irrelevant reciprocal influences were so important that children performed at random or even under chance level on incongruent trials whatever the task. These results thus confirm the presence of automatic numerical and perceptual processing in preschoolers as young as 4 years old.

E2 also provides further support to the presumption of different developmental trajectories for automatic numerical and perceptual processing. In the perceptual task, the influence of numerical over perceptual processing grew with age, a trend that was already observed in E1. By contrast, the congruity effect was not modulated by age in the numerical task. The irrelevant perceptual influences might even be greater in younger children, as 4-year-olds performed significantly under chance level on incongruent trials, whereas older age groups simply performed at random. This interpretation must nevertheless be taken cautiously given the absence of any significant interaction between congruity and age group in the numerical task.

The direct comparison between stimuli presented in E1 and E2 using the same experimental design in children from the same age group and preschool level also yielded interesting results. Perceptual comparison was not affected by the type of stimuli, but for the numerical task, dot stimuli yielded superior performance compared with bars. This result supports our initial intuition that the uncontrolled total circumference (and diameter) indirectly contributed to performance in the numerical task when comparing collections of dots, which led to overestimating the efficiency of numerosity processing in E1. In E2 where such confound was absent, performances in each task no longer differed from each other, attesting that numerical and perceptual tasks were of similar complexity when comparing collections of bars.

Finally, despite the use of stimuli enabling a better control of correlated perceptual variables, E2 also failed to find evidence for a specific relationship between the acquisition of counting knowledge and the development of numerosity processing under both intentional and unintentional conditions. As the type of stimuli presented can no longer be incriminated, an alternative explanation for these nonsignificant results could be the age of participants tested so far in the first two experiments. The acquisition of counting knowledge could indeed play a role in the development of numerosity processing only at the very beginning, when children come to understand how verbal numerical labels represent numerosities independently of perceptual variations. According to Wynn (1990, 1992) and Fuson (1992), this conceptual shift takes place around the age of 3. In keeping with this hypothesis, evidence showed that children who lacked minimal counting or cardinal knowledge were unable to perform true numerical comparison, for example, to recognize the numerical equivalence between sets of heterogeneous objects (Mix, 1999b) or to compare visual numerosities when surface area was controlled for (Brannon & Van de Walle, 2001; Rousselle et al., 2004). In Brannon and Van de Walle's (2001) study, the correlation between success in the numerical comparison and verbal counting skills even disappeared when children with no counting or cardinal knowledge were removed from the analysis, suggesting that the improvement of numerosity processing is only related to acquisition of the very first counting knowledge but not to the development of subsequent counting skills.

If this hypothesis is correct, we would expect a specific relationship between counting skills and numerosity processing to emerge in children from 3 years of age, who only start to grasp how counting words map onto numerosities. The last experiment was thus conducted to examine the development of automatic and intentional numerosity processing in 3-year-olds and to test the replicability of the results obtained in 4- and 5-year-olds using the same method as in E2.

Experiment 3

Method

Participants

Fifty-seven Caucasian preschoolers participated in this experiment. They were drawn from a local middle-class preschool in

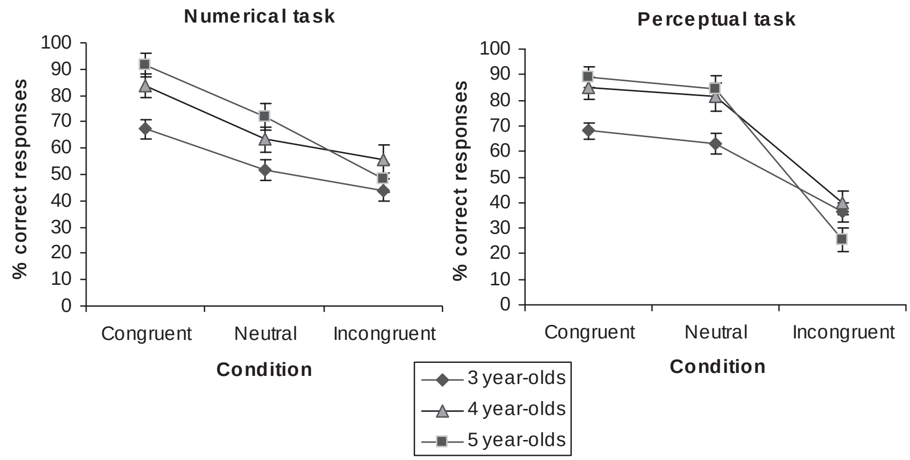

Figure 5. Experiment 3: Congruity effect in the numerical and perceptual comparison tasks; percentages of correct responses and standard errors for each age group.

Brussels, Belgium, and were of middle to upper socioeconomic status. Children were divided into three age groups: 3-year-olds (n = 25; mean age = 3 years, 7 months; SD = 3.2 months; range: 3 years, 0 months to 3 years, 11 months; 11 girls, 14 boys), 4-year-olds (n = 16; mean age = 4 years, 6 months; SD = 3.3 months; range: 4 years, 0 months to 4 years, 10 months; 10 girls, 6 boys), and 5-year-olds (n = 16; mean age = 5 years, 6 months; SD = 3.8 months; range: 5 years, 0 months to 5 years, 11 months; 11 girls, 5 boys). Three-, 4-, and 5-year-olds, respectively, attended lower, intermediate, and higher level preschool classes. No other information about participants was available.

Materials and Task Procedure

Tasks, stimuli, and procedure were the same as in E2, except that the testing took place in December. All children included in the present sample met the familiarization criterion in the Stroop collection comparison task. Nevertheless, it is worth noting that three 3-year-olds failed to meet this criterion and did not complete the test phase.

Results

Collection Comparison Stroop Tasks

For the same rationale as in E1 and E2, accuracy data in each tasks were analyzed in a 3 (condition: congruent, neutral, incongruent) × 3 (age group: 3-year-olds, 4-year-olds, 5-year-olds) repeated-measures MANOVA.

Figure 5 presents percentages of correct responses by task and congruity condition for each group (see also Table 7). In the numerical task, the main effects of condition and age group were significant, $F(2, 53) = 39.12$, $p < .001$, $\eta^2 = .60$, and $F(2, 54) = 7.98$, $p < .005$, $\eta^2 = .23$, respectively, but these two factors did not interact with each other, $F(4, 106) = 1.76$, $p > .10$. Pairwise comparison showed that all conditions differed significantly from one another (ps $< .001$) and that 3-year-olds performed lower than the two older groups (ps $< .05$), who had similar performance ($p > .10$).

In the perceptual task, the main effects and the interaction were significant: age group, $F(2, 54) = 10.08$, $p < .001$, $\eta^2 = .27$; condition, $F(2, 53) = 64.43$, $p < .001$, $\eta^2 = .71$; interaction, $F(4, 106) = 2.56$, $p < .05$, $\eta^2 = .09$. Three-year-olds had significantly lower performance than 4- and 5-year-olds (ps $< .01$), but the two older groups did not differ from each other ($p = 1$). There were significant congruity and interference effects (ps $< .001$) but no facilitation effect ($p > .10$). This condition effect was significant in all age groups and increased with age: mean congruity effect = 32%, 45%, and 64%, and $F(2, 23) = 11.71$, $F(2, 14) = 14.34$, and $F(2, 14) = 60.90$ in 3-, 4-, and 5-year-olds, respectively, all ps $<$ .001.[6]

Two-tailed t-tests were used to compare the percentages of correct responses for each age group with the percentage expected by chance (i.e., 50%) in each task and congruity condition. In the numerical task, all age groups performed above chance level on congruent trials: 3-year-olds, $t(24) = 4.47$; 4- and 5-year-olds, ts $(15) > 6.51$, all ps $< .001$. In the neutral condition, 3-year-olds performed at random, $t(24) = .45$, $p > .10$, whereas the two older groups performed significantly above chance level, ts $(15) > 2.26$, ps $< .05$, pointing to the great reliance on irrelevant perceptual dimensions in younger children. Finally, all groups performed at random on incongruent trials: 3-year-olds, $t(24) = -1.73$; 4-year-olds, $t(15) = 1.10$; 5-year-olds, $t(15) = -.23$, ps $> .10$. In the perceptual task, children from all age groups performed above chance level on congruent and neutral trials: 3-year-olds, ts $(24) >$ 2.61; 4- and 5-year-olds, ts $(15) > 6.64$, all ps $< .01$. Conversely, in the incongruent condition, all groups performed significantly under chance level, although this trend was only marginally significant in 4-year-olds: 3-year-olds, $t(24) = -3.55$, $p < .005$;

[6] A 2 × 3 MANOVA with order (numerical task first, perceptual task first) and age group (4-year-olds, 5-year-olds, 6-year-olds) as between-subjects factors and congruity effect as the dependent variable was carried out in each task to assess the effect of task order. In both tasks, neither the effect of order, $F(1, 51) < 1.48$, $p > .05$, nor the Order × Age group interaction, $F(2, 51) < 2.49$, $p > .05$, was significant.

Table 7
Experiment 3: Mean Performance in the Stroop Collection Comparison Task for Each Age Group

Task	3 years			4 years			5 years			Total		
	M	*SD*	CI (95%)	*M*	*SD*	CI (95%)	*M*	*SD*	CI (95%)	*M*	*SD*	CI (95%)
Perceptual Stroop task												
Congruent	68.0	17.9	61.3–74.7	84.8	17.8	76.4–93.1	89.1	13.0	80.7–97.4	78.6	19.0	73.6–83.7
Neutral	63.0	24.9	54.5–71.5	81.3	18.8	70.7–91.8	84.4	16.1	73.8–95.0	74.1	23.0	68.0–80.2
Incongruent	36.3	19.4	28.6–43.9	39.8	20.8	30.3–49.3	25.4	16.2	15.9–34.9	34.2	19.5	29.0–39.4
Total	55.8	11.1	51.8–59.7	68.6	8.1	63.7–73.6	66.3	9.4	61.3–71.2	62.3	11.4	59.3–65.3
Numerical Stroop task												
Congruent	67.3	19.3	60.1–74.4	83.6	20.7	74.6–92.6	91.8	11.6	82.8–100.8	78.7	20.6	73.3–84.2
Neutral	51.5	16.7	43.8–59.2	63.3	23.5	53.6–72.9	71.9	18.5	62.2–81.5	60.5	20.8	55.0–66.1
Incongruent	44.0	17.4	35.4–52.6	55.9	21.2	45.1–66.6	48.4	26.9	37.7–59.2	48.6	21.6	42.8–54.3
Total	54.3	11.1	48.6–59.9	67.6	17.3	60.5–74.7	70.7	15.0	63.6–77.8	62.6	15.8	58.4–66.8
Counting task												
Enumeration	4.1	3.3	2.4–5.7	10.4	5.3	8.3–12.4	13.3	3.8	11.2–15.3			
Trade	3.5	1.9	2.3–4.6	6.2	3.9	4.8–7.6	9.6	2.8	8.2–11.1			
Point–to–*x*	3.1	1.7	2.0–4.2	4.8	1.7	3.4–6.1	7.1	4.4	5.8–8.5			

Note. CI = confidence interval.

4-year-olds, $t(15) = -1.96$, $p = .07$; 5-year-olds, $t(15) = -6.07$, $p < .001$.

Counting Tasks

Table 8 presents Pearson correlations between age, counting scores (see Table 7), percentages of correct responses for neutral stimuli in the numerical task, and the interference effect in the perceptual task. Correlations with age were all significant. Moreover, the scores from the three counting tasks correlated significantly with each other as well as with performance in the neutral condition of the numerical task and with the interference effects in the perceptual task. However, when age was partialled out, only the correlations between enumeration and trade scores and between enumeration scores and performance in the neutral condition remained significant. Note that running the same analyses in 3-year-olds yielded no significant correlation, whether or not age was controlled for (all *p*s > .05).

Discussion

E3 replicates and extends the results obtained in E2. In all age groups, the irrelevant dimension affected performance in the numerical and perceptual tasks to such an extent that children performed at random and under chance level, respectively, on incongruent trials. Furthermore, the influence of numerical over perceptual processing increased with age, whereas the influence of perceptual over numerical processing was not modulated by age group.

A new result in E3 is that 3-year-olds performed at random on neutral trials in the numerical task, whereas 4- and 5-year-olds performed significantly above chance level as in E1 and E2. This corroborates previous findings showing that 3-year-olds were unable to select the larger of two collections of bars when surface area was controlled for (Rousselle et al., 2004) and hint at a greater influence of irrelevant perceptual processing in younger children. However, how could numerosity processing run automatically in the perceptual task and simultaneously be unable to operate intentionally in the numerical task? A possible way to reconcile these contrasting observations would be to consider that numerosity representations are accessible through different levels of consciousness: Numerosity information would be processed at a very low level of consciousness in the perceptual task, whereas processing numerosity under intentional conditions would tap into higher levels of consciousness that are still not reached by 3-year-olds. In this respect, Karmiloff-Smith (1997) speculated that infants' first knowledge about the world is implicit and, thus, not accessible to conscious introspection. As infants interact with the environment, this knowledge would undergo a gradual process of *representational redescription,* which involves recoding representations from an implicit format to an increasingly explicit format (i.e., quasi-linguistic representation) accessible to consciousness.

Supporting this hypothesis, several studies have demonstrated the existence of a discontinuity between infants' and preschoolers' competence. Although infants discriminate the numerosity of heterogeneous (Starkey et al., 1990; Strauss & Curtis, 1981) and sequential objects (Canfield & Smith, 1996; Wynn, 1996), match the numerosity of stimuli across different sensory modalities (Starkey et al., 1990), and perform simple arithmetic operations on visual stimuli (Wynn, 1992), 3-year-olds, in contrast, were found to be unable to recognize the numerical equivalence between sets of heterogeneous (Mix, 1999b), sequential (Mix, 1999a), or cross-modal stimuli (Mix et al., 1996) or to compute simple additions and subtractions (Houdé, 1997; Huttenlocher, Jordan, & Levine, 1994; Vilette, 2002). These discrepancies suggest that implicit paradigms are sometimes more susceptible to eliciting a particular process than explicit tasks and might explain why 3-year-olds are influenced by numerical properties in the perceptual task before being able to process numerosities intentionally in the numerical task.

Finally, E3 yielded uncertain results regarding the specific relationship between counting skills and numerosity processing, which was expected to emerge by testing younger children. Indeed, only the correlation between general counting performance (enumeration task) and intentional numerosity processing (neutral

Table 8
Experiment 3: Correlations Between Counting Tasks, Performance for Neutral Stimuli in the Numerical Task, and Interference Effect in the Perceptual Task

Variable	1	2	3	4	5	6
1. Age	—					
2. Enumeration	.70***	—	.45***	.19	.33*	−.05
3. Trade	.72***	.72***	—	.11	.11	.13
4. Point-to-*x*	.52***	.48***	.44***	—	.01	−.07
5. N task: Neutral trials	.45***	.52***	.39**	.24†	—	−.17
6. P task: Interference	.43***	.26*	.39**	.17*	.06	—

Note. The area below the diagonal presents Pearson correlation coefficients and the above the diagonal reports partial correlations controlling for age. N task = numerical task; P task = perceptual task.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

trials of the numerical task) remained significant when age variations were controlled for. As counting performance in the enumeration task is determined by multiple processing subcomponents (pointing, reciting the number word sequence and applying a one-to-one correspondence between recitation and pointing), it is difficult to establish which subcomponent actually plays a role in the numerical development, all the more so because most of them probably also contributed to performance in the trade and point-to-*x* tasks, which, in contrast, did not correlate with numerosity processing when age was controlled for.

In E2, we speculated that the improvement of numerosity processing could be induced by acquisition of the first counting knowledge. Regarding this hypothesis, one possibility remains that most 3-year-olds tested in E3 already passed the critical period of grasping the relationship between counting words and numerosities and, thus, already possessed the decisive and minimal counting or cardinal knowledge by the time they were tested here. In support of this assumption, only 1 child never gave a single cardinal response (i.e., 1.7%), and 2 provided a cardinal response only when required to point to the collection including a single object (i.e., 3.5%); that is, a total of 5.2% of our total sample did not have minimal cardinal knowledge in E3. This proportion contrasts with the larger number of children lacking minimal cardinal knowledge in studies showing a relationship between numerical performance and counting development, that is, 34% in Brannon and Van de Walle's (2001) study, 43–54% in Rousselle et al.'s (2004) study, and 17–47% in Mix's (1999b) study. If the improvement of numerosity processing is actually related to acquisition of the very first counting knowledge but not to the development of subsequent counting skills, then the absence of a specific relationship between counting and numerosity processing development might be related to the very small number of children who still lacked minimal cardinal knowledge in the present sample, in contrast to other studies.

General Discussion

The aim of this study was first to determine whether preschoolers process numerical and perceptual dimensions automatically in continuous and discrete quantification tasks. A second and related issue was to examine how the automatic processing of irrelevant numerical and perceptual properties is modulated throughout development with age as well as the gradual acquisition of counting skills and the growing maturity of inhibition processes.

Using a Stroop paradigm, the present three experiments showed significant and reciprocal interference of the irrelevant numerical and perceptual dimensions in the required processing in preschoolers. Moreover, in both tasks, incongruent trials invariably led to drastic performance decrements in all age groups. These results provide converging evidence for the mandatory processing of numerical and perceptual magnitude during the quantification of nonsymbolic stimuli in children beginning at age 3.

Furthermore, all three experiments consistently found that the congruity effect size grew with age in the perceptual task but not in the numerical task. This evidence suggests that the influence of numerosity over perceptual quantification increased with age despite the growing efficiency of inhibition processes, whereas the influence of perceptual cues over numerical quantification tended rather to remain stable (or even to decrease) during childhood. In this respect, it can be argued that the modulation of the congruity effect across age in the perceptual task is constrained by ceiling effects in the youngest age groups. Indeed, 3- and 4-year-olds' mean performance in E2 and E3 did not exceed 70% in both tasks, indicating that processing bar stimuli was difficult for them. However, the presence of ceiling effects in both tasks does not explain why the congruity effect grew with age in the perceptual task only.

Many authors have argued that discrete and continuous quantities are represented through an innate single magnitude system (Feigenson, 2007; Gallistel & Gelman, 1992, 2000; Walsh, 2003). However, despite the existence of a reciprocal relationship between numerosity and perceptual processing in the early quantification task, the presence of a divergent developmental course in their automatization challenges the hypothesis of a common processing mechanism operating from birth. Rather, the present pattern of results suggests that the automatic access to perceptual information has already reached maturity in preschoolers, whereas the automatization of numerosity processing is only achieved gradually during childhood, as with the automaticity in accessing number magnitude from Arabic symbols (Girelli et al., 2000; Noël, Rousselle, & Mussolin, 2005; Rubinsten et al., 2002).

The present findings thus provide support to the hypothesis of a gradual extraction of numerosity, becoming an increasingly salient dimension of the environment (Mix et al., 2002). In line with previous studies showing that infants and even preschoolers do not discriminate between numerosities when correlated perceptual variables were strictly controlled for (Clearfield & Mix, 1999, 2001; Feigenson et al., 2002; Rousselle et al., 2004), our results suggest that numerical processing takes longer to be fully operational, compared with perceptual processing. However, although developmental changes observed in the numerical and perceptual tasks undoubtedly reflect the progressive transformations of the underlying representational systems across age, we can never exclude the possibility that other factors might have also played a role in these changes. For example, children's comprehension and memory of the instructions in either task or even their beliefs about what we expected them to do might also vary with age and therefore might have influenced the pattern of results.

Besides the issues addressed in this study, the use of the same experimental design in children from the same age group and preschool level in E1 and E2 provided a unique opportunity to assess how the pattern of performance is affected by the type of stimuli presented. Dot stimuli yielded superior performance compared with bars in the numerical task, whereas perceptual comparison was unaffected by the type of stimuli. This result attests for the contribution of an uncontrolled perceptual variable (total circumference and diameter) to performance in the numerical task when comparing collections of dots and adds to the growing body of evidence that casts doubts on the nature of the processes underlying infant's visual quantification (Clearfield & Mix, 1999, 2001; Feigenson et al., 2002; Rousselle et al., 2004). Most studies on infant's numerical cognition used collections of dots to demonstrate the presence of number-based discrimination behavior. The present results prove that the use of such stimuli led to overestimating the efficiency of visual numerosity processing, as this kind of stimuli does not allow simultaneous control of the summed surface area and the summed circumference (and summed diameter) of visual arrays. On the other hand, the single elements might be difficult to individuate for our visuospatial attention system in collections of vertical bars packed together. Nevertheless, the need for applying efficient perceptual controls in studies on numerosity processing development is now clearly demonstrated.

Some particular counting requirements inherent to the abstraction, as well as the one-to-one correspondence and the cardinality principles, led us to assume that their gradual mastery would have contributed to the development of numerosity processing under both intentional and unintentional conditions. Unfortunately, correlational analyses yielded unconvincing results in all three experiments. Although the improvement of counting generally correlated with intentional and unintentional numerosity processing development, it was not possible to determine whether this relationship was due to counting improvement per se or to some age-related changes. In E2, we speculated that the acquisition of counting knowledge could play a role in numerosity processing development only at the very beginning, when children start to grasp how counting words represent numerosities in the real world. However, testing younger children in E3 did not help find strong evidence for a specific relationship between counting and numerosity processing development, as correlations disappeared when age was partialled out.

These results are inconsistent with previous reports (Brannon & Van de Walle, 2001; Mix, 1999a, 1999b; Mix et al., 1996, 1997; Rousselle et al., 2004). One possibility for this is that the children in our sample, even the younger ones, already had a "too-good" level of counting to allow the observation of a significant relationship with the processing of numerosities. A second possibility is that the counting tasks were unreliable measurements of children's actual level of counting development. Children might have been discouraged by the necessity of counting many times in succession and thus might have just "botched" some tasks. However, performance in counting tasks strongly correlated with each other and with age, which is inconsistent with this hypothesis. Another possibility is that children mainly used procedural knowledge to solve the counting tasks, including the point-to-*x* task, which was supposed to assess the cardinality principle development. They might have used the counting procedure to find the card depicting a particular number of items, without really understanding the cardinal meaning of number words. This hypothesis also seems unlikely. Indeed, although the reliability of the trade and enumeration tasks can be questioned, it has been shown that the point-to-*x* task is a reliable measurement of cardinal concept development (Wynn, 1992).

For now, the present study suggests that the development of primitive numerosity processing was unrelated to the acquisition of counting skills. One plausible interpretation is that learning to count might be contributive to the formation of exact numerical representations (Carey, 2001) but not to the development of approximate number magnitude representations involved in the processing of large numerosities, as in the present comparison tasks. In support of this hypothesis, the relationship between counting and numerosity processing development has been almost exclusively observed with tasks involving small numerosities (Brannon & Van de Walle, 2001; Mix, 1999a, 1999b; Mix et al., 1996, 1997). The only study that reported no relationship between verbal counting and performance in a numerical comparison task involved large numerosities (Huntley-Fenner & Cannon, 2000).

Nevertheless, to confirm this hypothesis, the present results should be replicated using other indicators of cardinality conceptual development, such as the give-a-number task, whose reliability is now well established in the literature (Brannon & Van de Walle, 2001; Mix, 1999a, 1999b; Mix et al., 1996, 1997; Rousselle et al., 2004; Sarnecka & Gelman, 2004; Wynn, 1992), or the compare-sets task, designed to test whether children know that number words refer to specific numerosities (Sarnecka & Gelman, 2004). Likewise, the spontaneous use of one-to-one correspondence to produce sets of equivalent number has to be evaluated through other situations before concluding that it has no impact on numerosity processing development. Other aspects of counting development could also be investigated to clarify the link between symbolic and nonsymbolic numerical processing development. For instance, the mastery of the order-irrelevance principle (i.e., counting from the beginning or from the end does not change the numerosity) might help children understand the independence between the collection layout and its numerosity.

To summarize, this study traced the development of automatic numerical and perceptual processing of collections in children with no formal school experience and little symbolic numerical knowledge. Results show that numerosities start to interfere with perceptual processing at age 3 and that the sensitivity to irrelevant numerical cues increases with age despite children's growing ability to resist to interference. Reciprocally, perceptual properties start to interfere with the processing of numerosity at the same age, but these influences remain quite stable throughout development in preschool. These findings support the idea that the automatization of numerosity processing arises gradually over the course of the development, whereas the automatic access to perceptual information is already well developed in preschoolers. The kind of stimuli presented has a great influence on the pattern of performance in numerical tasks, and using dots stimuli clearly leads to overestimate the efficiency of visual numerosity processing. The question of whether or not the acquisition of some counting knowledge specifically contributed to the development of numerosity processing still remains to be clarified.

References

Antell, S. E., & Keating, D. P. (1983). Perception numerical invariance in neonates. *Child Development, 54,* 695–701.

Bijeljac-Babic, R., Bertoncini, J., & Melher, J. (1993). How do four-day-old infants categorize multisyllabic utterances? *Developmental Psychology, 29,* 711–721.

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

Brannon, E. M., & Van de Walle, G. A. (2001). The development of ordinal numerical competence in young children. *Cognitive Psychology, 43,* 53–81.

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

Canfield, R. L., & Smith, E. G. (1996). Number-based expectations and sequential enumeration by 5-month-old infants. *Developmental Psychology, 32,* 269–279.

Carey, S. (2001). Cognitive foundations of arithmetic: Evolution and ontogenesis. *Mind & Language, 16,* 37–55.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72,* 1032–53.

Clearfield, M. W., & Mix, K. S. (1999). Number vs contour length in infants' discrimination of small visual sets. *Psychological Science, 10,* 408–411.

Clearfield, M. W., & Mix, K. S. (2001). Amount versus number: Infants' use of area and contour length to discriminate small sets, *Journal of Cognition and Development, 2,* 243–260.

Dehaene, S., & Akhavein, R. (1995). Attention, automaticity and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21,* 314–326.

Dehaene, S., & Changeux, J-P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience, 5,* 390–407.

den Heyer, K., & Briand, K. (1986). Priming single digit numbers: Automatic spreading activation dissipates as a function of semantic distance. *American Journal of Psychology, 99,* 315–340.

Dormal, V., Seron, X., & Pesenti, M. (2006). Numerosity-duration interference: A Stroop experiment. *Acta Psychologica, 121,* 109–124.

Droit-Volet, S., Clément, A., & Fayol, M. (2003). Time and number discrimination in a bisection task with a sequence of stimuli: A developmental approach. *Journal of Experimental Child Psychology, 84,* 63–76.

Duncan, E. M., & McFarland, C. E. (1980). Isolating the effects of symbolic distance and semantic congruity in comparative judgments: An additive-factors analysis. *Memory & Cognition, 8,* 612–622.

Feigenson, L. (2007). The equality of quantity. *Trends in Cognitive Science, 11,* 185–187.

Feigenson, L., Carey, S., & Hauser, M. (2002). The representation underlying infants' choice of more: Object files vs. analog magnitudes. *Psychological Science, 13,* 150–156.

Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology, 44,* 33–66.

Fuson, K. S. (1992). Relationships between counting and cardinality from age 2 to age 8. In J. Bideaud, C. Meljac, & J-P. Fischer (Eds.), *Pathways to number: Children's developing numerical abilities* (pp. 127–149). Hillsdale, NJ: Erlbaum.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44,* 43–74.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences, 4,* 59–65.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3.5–7 years old on a Stroop-like day–night test. *Cognition, 53,* 129–153.

Girelli, L., Lucangeli, D., & Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *Journal of Experimental Child Psychology, 76,* 104–122.

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175–204). New York: Academic Press.

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition, 10,* 389–395.

Houdé, O. (1997). Numerical development: From the infant to the child. Wynn's (1992) paradigm in 2- and 3-year olds. *Cognitive Development, 12,* 373–391.

Howell, D. C. (2001). *Statistical methods for psychology* (5th ed.). Belmont, CA: Duxbury Press.

Huntley-Fenner, G., & Cannon, E. (2000). Preschoolers' magnitude comparisons are mediated by a preverbal analog mechanism. *Psychological Science, 11,* 147–152.

Huttenlocher, J., Jordan, N., & Levine, S. (1994). A mental model for early arithmetic. *Journal of Experimental Psychology: General, 123,* 284–296.

Karmiloff-Smith, A. (1997). *Beyond modularity: A developmental perspective on cognitive science* (3rd ed.). Cambridge, MA: MIT Press.

Lipton, J. S., & Spelke, E. S. (2003). Origins of the number sense: Large-number discrimination in human infants. *Psychological Science, 14,* 396–401.

Lipton, J. S., & Spelke, E. S. (2004). Discrimination of large and small numerosities by human infants. *Infancy, 5,* 271–290.

Mix, K. S. (1999a). Preschoolers' recognition of numerical equivalence: Sequential sets. *Journal of Experimental Child Psychology, 74,* 309–332.

Mix, K. S. (1999b). Similarity and numerical equivalence: Appearances count. *Cognitive Development, 14,* 269–297.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (1996). Do preschool children recognize auditory-visual numerical correspondences? *Child Development, 67,* 1592–1608.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). Multiple cues for quantification in infancy: Is number one of them? *Psychological Bulletin, 128,* 278–294.

Mix, K. S., Levine, S. C., & Huttenlocher, J. (1997). Numerical abstraction in infants: Another look. *Developmental Psychology, 33,* 423–428.

Morin, R. E., DeRosa, D. V., & Stultz, V. (1967). Recognition memory and reaction time. *Acta Psychologica, 27,* 298–305.

Noël, M-P., Rousselle, L., & Mussolin, C. (2005). Magnitude representation in children: Its development and dysfunction. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition.* New York: Psychology Press.

Piaget, J., & Szeminska, A. (1964). *La genèse du nombre chez l'enfant* [The child's conception of number] (3rd ed.). Neuchâtel, Switzerland: Delachaux et Niestlé.

Rousselle, L., Palmers, E., & Noël, M-P. (2004). Magnitude comparison in preschoolers: What counts? Influence of perceptual variables. *Journal of Experimental Child Psychology, 87,* 57–84.

Rubinsten, O., Henik, A., Berger, A., & Shahar-Shalev, S. (2002). The development of internal representations of magnitude and their association with Arabic numerals. *Journal of Experimental Child Psychology, 81,* 74–92.

Sarnecka, B. W., & Gelman, S. A. (2004). Six does not just mean a lot: Preschoolers see number words as specific. *Cognition, 92,* 329–352.
Simon, T. J. (1997). Reconceptualizing the origins of number knowledge: A "non numerical" account. *Cognitive Development, 12,* 349–372.
Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition, 36,* 97–127.
Strauss, M. S., & Curtis, L. E. (1981). Infant perception of numerosity. *Child Development, 52,* 1146–1152.
Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101,* 80–102.
Tzelgov, J., & Ganor-Stern, D. (2005). Automaticity in processing ordinal information. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition.* New York: Psychology Press.
Tzelgov, J., Meyer, J., & Henik, A. (1992). Automatic and intentional processing of numerical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 166–179.
Vilette, B. (2002). Do young children grasp the inverse relationship between addition and subtraction? Evidence against early arithmetic. *Cognitive Development, 17,* 1365–1383.
Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences, 7,* 483–488.
Wood, J. N., & Spelke, E. S. (2005a). Chronometric studies of numerical cognition in five-month-old infants. *Cognition, 97,* 23–39.
Wood, J. N., & Spelke, E. S. (2005b). Infant's enumeration of actions: Numerical discrimination and its signature limits. *Developmental Science, 8,* 173–181.
Wynn, K. (1990). Children's understanding of counting. *Cognition, 36,* 155–193.
Wynn, K. (1992, August 27). Addition and subtraction by human infants. *Nature, 358,* 749–750.
Wynn, K. (1995). Origins of numerical knowledge. *Mathematical Cognition, 1,* 35–60.
Wynn, K. (1996). Infants' individuation and enumeration of actions. *Psychological Science, 7,* 164–169.
Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations, *Cognition, 89,* B15–B25.
Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74,* B1–B11.
Xu, F., & Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science, 8,* 88–101.
Zbrodoff, N. J., & Logan, G. D. (1986). On the autonomy of mental processes: A case study of arithmetic. *Journal of Experimental Psychology: General, 115,* 118–130.

Received July 21, 2006
Revision received August 1, 2007
Accepted September 5, 2007 ■