Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
J. Experimental Child Psychology 87 (2004) 57–84
Journal of Experimental Child Psychology
www.elsevier.com/locate/jecp

Magnitude comparison in preschoolers: What counts? Influence of perceptual variables

Laurence Rousselle,[a] Emmanuelle Palmers,[a] and Marie-Pascale Noël[a,b,*]

[a] *Unité de Neurosciences Cognitives, University of Louvain, 1348 Louvain-la-Neuve, Belgium*
[b] *Unité Cognition et Développement, Faculté de Psychologie, Université Catholique de Louvain, 10 Place C. Mercier, 1348 Louvain-la-Neuve, Belgium*

Received 4 August 2003; revised 20 October 2003

Abstract

This study examined numerosity comparison in 3-year-old children. Predictions derived from the analog numerical model and the object-file model were contrasted by testing the effects of size and ratio between numerosities to be compared. Different perceptual controls were also introduced to evaluate the hypothesis that comparison by preschoolers is based on correlated perceptual variables rather than on number per se. Finally, the relation between comparison performance and verbal counting knowledge was investigated. Results showed no evidence that preschoolers use an analog number magnitude or an object-file mechanism to compare numerosities. Rather, their inability to compare sets controlled for surface area suggests that they rely on perceptual cues. Furthermore, the development of numerosity-based representations seems to be related to some understanding of the cardinality concept.

Keywords: Children; Development; Numerosities; Number; Counting; Magnitude; Comparison

Introduction

Over the past 20 years, habituation experiments have provided evidence that young infants, even neonates, can discriminate the numerosity of small collections of up to three or four elements, whether these are simultaneous (Antell & Keating,

* Corresponding author. Fax: +32-10-47-37-74.
E-mail address: Marie-pascale.noel@psp.ucl.ac.be (M.-P. Noël).

0022-0965/$ - see front matter
doi:10.1016/j.jecp.2003.10.005

1983; Starkey & Cooper, 1980; Strauss & Curtis, 1981), sequential (Wynn, 1996), or moving visual stimuli (van Loosbroek & Smitsman, 1990) or even auditory stimuli (Bijeljac-Babic, Bertoncini, & Mehler, 1993). Other studies using the methodology of expectancy violation have shown that infants can anticipate the result of small additions and subtractions (Wynn, 1992a; Simon, Hespos, & Rochat, 1995; Koechlin, Dehaene, & Mehler, 1997). Furthermore, infants seem to be able to detect numerical correspondence between sets of entities presented in different sensory modalities (Starkey, Spelke, & Gelman, 1990; but see Moore, Benenson, Reznick, Peterson, & Kagan, 1987; Mix, Levine, & Huttenlocher, 1997). These results suggest that infants are able to operate at a remarkably abstract level, a level that could serve as a starting point for numerical reasoning.

Theories of early quantification

These exciting results have led many researchers to assume that infants possess an innate mechanism for representing numerosity. Two numerical models have been proposed that both assume an analog representation of quantity: the neuronal model (Dehaene & Changeux, 1993) and the accumulator model (Gallistel & Gelman, 1992). As their basic assumptions and their predictions are quite similar, only the latter is described in detail. The accumulator model was initially proposed by Meck and Church (1983) to account for animals' competencies and was later applied to infants by Gallistel and Gelman (1992) and Wynn (1995). This model supposes a preverbal counting mechanism that works as follows. A pacemaker emits pulses of energy at a constant rate. Each time an entity is counted, a switch closes for a brief and fixed temporal interval, allowing energy to pass into an accumulator. For each entity counted, the accumulator fills up in equal-sized increments. The resulting state of the accumulator is an analog representation of the number of entities counted. However, this representation has a precision that is inversely proportional to the numerosity considered (Weber's law). This variability can account for the numerical distance and size effects or the numerical ratio effect, which combines the two others. The numerical distance effect refers to the finding that the ability to discriminate two numbers improves as the numerical distance between them increases (e.g., it is easier to compare 4 with 9 than 4 with 5). The number size effect refers to the finding that, for equal numerical distance, discrimination of two numbers declines as their numerical size increases (e.g., it is more difficult to compare 8 with 9 than 2 with 3). The ratio effect corresponds to the decrease in the numerical discrimination when the ratio between the two numerosities approaches 1 (e.g., it is easier to compare 2 with 4 (ratio of 1/2) than 8 with 10 (ratio of 4/5)). This effect combines the two others because ratio covaries simultaneously with the size and the distance of the two compared numerosities. Indeed, on the one hand, for equal numerical distance, the ratio approaches 1 when the numerosities increase (4:8, 8:12, 12:16, 16:20, ...). On the other hand, the ratio approaches 1 when the distance between the two numerosities decreases (1:10, 1:8, 1:6, 1:4, ...).

For other researchers, however (Simon, 1997; Uller, Carey, Huntley-Fenner, & Klatt, 1999), numerical tasks are successfully performed thanks to general nonnumerical competences. In their model, designated here by the "object-file" model, they speculate that infants' behavior is a by-product of a preattentive individuation process for keeping track of visible objects (Kahneman, Treisman, & Gibbs, 1992; Trick & Pylyshyn, 1994). This process, which has been evoked to explain subitizing in adults (i.e., accurate and almost immediate apprehension of small numerosities), would account for the fact that infants discriminate between sets up to three or four items. In this object-tracking mechanism, a reference token is assigned to each distinct feature cluster in a scene. These reference tokens are limited in number (up to four in general) and are assigned in parallel. They are used to code the object's location as long as it remains visible. The object-file model assumes that these tokens are used by infants and young children to respond to quantitative tasks. Faced with a set of objects, they would represent each object by one of these tokens, which can be held in memory when the set disappears. Numerosity discrimination would result from the observation of a mismatch in the one-to-one correspondence between mentally stored reference tokens and objects of the new set. There would thus be no representation of numerosity per se.

In this view, infants would be able to discriminate only numerosities that can be handled by this object-file mechanism (i.e., up to 3 or 4 elements). A size effect is thus predicted, as only small (and not medium or large) collections could be discriminated. However, several authors (Xu & Spelke, 2000; Xu, 2003; Lipton & Spelke, 2003) found that 6-month-old infants can discriminate large numerosities if they differ by a ratio of 1:2. For example, infants discriminate 8 versus 16 elements but not 8 versus 12. A ratio effect on large numerosities has also been reported by Huntley-Fenner and Cannon (2000) in 3- to 5-year-old children: their performance on a magnitude comparison task was more accurate for a ratio of 1:2 (numerosities 1 vs 2 to 5 vs 10) than for a ratio of 2:3 (numerosities 2 vs 3 to 10 vs 15). The existence of a ratio effect supports the analog models of magnitude representation rather than the object-file model.

Recently, Mix, Huttenlocher, and Levine (2002) considered a third hypothesis to explain infants' quantification performance. They noticed some methodological bias in studies of infants related to a lack of control of perceptual cues that naturally covary with number, such as surface area, volume, contour length, and density. Because they roughly covary with number, perceptual and numerical quantification processes tend to yield the same reaction to a given situation. Thus, infants' reactions in a quantitative task may be guided by the detection of a change in number and/or by the perception of various continuous variations. In other words, when infants discriminate two from three objects, is it because both sets are numerically different or because three objects take more space than two? Among studies presenting visual stimuli simultaneously, most did not control for the surface area (Starkey & Cooper, 1980; Antell & Keating, 1983; Strauss & Curtis, 1981, 1984; Starkey et al., 1990; van Loosbroek & Smitsman, 1990; Cooper, 1984; Huntley-Fenner & Cannon, 2000; Wynn, 1992a; Simon et al., 1995).

Xu and Spelke (2000) elaborated a procedure to control this variable as well as brightness, array size, element size, and density by equating, in test displays, the cues that varied across habituation displays and by varying those that were equated.

However, they failed to control for the sum of contour length (or the diameter) of the elements, which remained confounded with numerosity (see Mix et al., 2002). The same criticism can be addressed to a study by Brannon (2002) in which she tested infants' ability to apprehend "greater than" and "less than" relations between numerosities. Infants were habituated to ascending or descending sequences of three numerical displays and then tested with both ascending and descending sequences (i.e., 4–8–16 or 16–8–4). Although the procedure controlled for total surface area, the sum of the contour length still covaried with number in two of three of the habituation sequences and was totally confounded with number in the test sequences.

Recently, Xu (2003) tried to overcome these limitations by using two conditions of perceptual control: one controlled for total filled area and the other for the sum of the contour length. She showed that 6-month-old infants discriminate four from eight elements in both perceptual control conditions. Yet, because area and contour length were never controlled simultaneously, this procedure allowed infants to use the cue that was not controlled for in each condition to quantify the set (total filled area in the contour length condition and contour length in the area condition).

Studies with sequential stimuli are also limited by methodological constraints related to a lack of control of rate, duration, and rhythmic pattern (Wynn, 1996; Canfield & Smith, 1996; Bijeljac-Babic et al., 1993; Starkey et al., 1990). Lipton and Spelke (2003) attempted to control the continuous temporal variables that covary with numerosity of auditory sequences by following the same procedure as Xu and Spelke (2000). The total duration and the amount of sound/silence (acoustic energy) varied during habituation and were equated during test. Moreover, the rate of individual sounds varied in habituation and test trials (during test, sequences of 16 sounds were twice as fast as sequences of 8). Results showed that 6-month-old infants exhibit a head-turn preference toward the novel numerosity. However, the total duration and the rate of the sequence were not completely dissociated in their procedure. Indeed, infants who were habituated to 8 sounds always heard slow–long and fast–short sequences. At test, the 8-sound sequences were slow and long as during habituation, but the 16-sound sequences were fast and long, which is a conjunction of auditory properties that the infants had never heard before. Conversely, infants who were habituated to 16 sounds always heard slow–long and fast–short sequences. In test, the 16-sound sequences were fast and short, as during habituation, but 8-sound sequences were slow and short, which is a conjunction of temporal properties that the infants had never heard during habituation.

After having examined all the results supporting the idea of early numerical competences in infants, Mix et al. (2002) concluded that there is no clear-cut evidence that infants' performance relies on number in quantitative tasks. Instead, they developed the hypothesis that infants represent discrete quantities in terms of overall amount, in a continuous way (amount of area, volume, length, etc.). Supporting this "amount" model, two studies that strongly controlled for perceptual variables failed to find any evidence that infants respond to numerosities. In a 1 vs 2 comparison, Feigenson, Carey, and Spelke (2002) found that when number and total front surface area are pitted against each other, infants dishabituated to a change in front surface area (i.e., habituation, one large object; test, one small object) but not to a change in number (i.e., habituation, one large object; test, two small objects). The same results were found with the methodology of violation of expectancy in addition task (1 small object + 1 small object = 1 large object or 2 large objects) and subtraction task (2 large objects − 1 large object = 1 small object or 2 small objects). Similarly, Clearfield and Mix (1999, 2001) observed that infants reacted to a change in contour length (i.e., sum of each individual item's perimeters of the display, which covaries with surface area) but not to a change in number (2 vs 3). These last two findings suggest that infants are more focused on the overall amount than on the number of objects. Given these contradictory results, we cannot at present conclude that infants are able to represent discrete number properties separately from correlated perceptual variables.

The first aim of the present study was to compare the three main propositions existing in the literature and to determine the type of representation really used by young children in quantitative tasks. When faced with quantities, do infants use a numerical preverbal counting mechanism specifically dedicated to number processing (the analog numerical model)? Do they exploit general nonnumerical competences based on the capacity to assign object tokens to each entity and to match them one by one (the object-file model)? Or do they perceive quantities in terms of overall amount, quantifying sets by using nonnumerical cues such as area, volume, and contour length (the amount model)?

These competing theories were tested in a comparison task. Children were presented with two collections of sticks and were asked to indicate which set has more. Ratio between pair members and size of the pairs were varied systematically to assess the independent effects of the two variables. Furthermore, different perceptual controls were introduced in the stimuli. In addition to the classically controlled density and contour (length and width) of the collections, some stimuli were also controlled for surface area of the sticks in the collections.

If children possess a numerical mechanism for representing numerosity, the analog numerical model predicts a ratio and possibly a size effect. Moreover, all perceptual control conditions should be performed at least above chance. If children use an object-file mechanism, they would be able to compare small numerosities only (size effect) and this even in the most controlled perceptual condition. Finally, if children do not process numerosity but base their judgment on continuous perceptual variables (the amount model), their performance should be less accurate (or even at chance) in the most controlled perceptual condition. Furthermore, in the least controlled condition in which most perceptual properties covary with number, they should be sensitive to the amount ratio between collections, which is then confounded with the numerical ratio.

These predictions should nevertheless be tempered. Indeed, the existence of a numerical or of an object-file mechanism does not preclude the possibility of processing and comparing amount. Therefore, an effect of perceptual conditions is also expected under these two theories: comparison might be easier in the condition in which most perceptual variables covary with number than in the one in which number is the only relevant cue. Yet, even in the most controlled perceptual condition, the numerical model predicts that performance should be above chance for large ratios (1/2), whereas the object-file model predicts that performance should be above chance for small numerosities.

The second question addressed in this paper concerns the relation between the acquisition of conventional counting system and the development of a number concept. Several studies present results suggesting the existence of such a relation whereas others do not. Mix et al. (Mix, 1999a, 1999b; Mix, Huttenlocher, & Levine, 1996; Mix et al., 1997) showed that more proficient counters performed more accurately than less proficient counters in numerical equivalence tasks. Moreover, although equivalence judgments between similar sets of visual stimuli were successfully performed by all children, only those who mastered the verbal counting system could make equivalence judgments between sets of sequential, heterogeneous, or cross-modal stimuli. Brannon and Van de Walle (2001) adopted an intermediate position according to which minimal verbal numerical competences are necessary to make quantitative judgments but quantitative competence is unrelated to further counting knowledge. Finally, Huntley-Fenner and Cannon (2000) found that the performance of 3- to 5-year-olds in numerosity comparisons of large collections was not predicted by their verbal counting ability.

A possible way to reconcile those contrasting results might be to differentiate between the situations in which a true numerical comparison is needed and those in which the comparison can be based on nonnumerical dimensions such as, for instance, the surface of the stimuli. If perceptual variables are not strictly controlled, tasks using sets of similar visual stimuli might be realized on nonnumerical grounds (e.g., the surface occupied by the stimuli) so that no counting abilities would be required (Huntley-Fenner & Cannon, 2000; Mix et al., 1996). When, in contrast, the task uses heterogeneous stimuli (Mix, 1999a) or manipulates the congruency of the perceptual variables with numerosity (Brannon & Van de Walle, 2001), then true numerical processing should take place. To perform those tasks correctly, the child has to be able to consider that each object counts for one entity, whatever its size or other physical properties, and that the comparison has to be based on that "entity" property and on nothing else. In that sense, counting practice (that is, counting collections rather than just reciting the number names) probably contributes to the development of that capacity. Accordingly, Mix (1999a, 1999b) as well as Brannon and Van de Walle (2001) found that only children with some mastery of counting could realize those types of tasks.

Depending on the theoretical position adopted, different predictions can be made concerning the relation between counting development and children's performance in quantification tasks. Gallistel and Gelman (1992) assume the existence of an innate preverbal numerical mechanism which works in accordance with counting principles. These nonverbal representations would precede and guide the acquisition of the verbal counting system. In this view, children should be able to do numerical comparison without any verbal counting mastery and their performance in comparison should be independent of their counting level. Conversely, if children do not possess a mechanism sensitive to discrete number properties as in the object-file or in the amount model, the acquisition of counting could play a role in the emergence of discrete quantification processes (Mix, 1999a, 1999b; Mix et al., 1996, 1997; Brannon &

Van de Walle, 2001). In particular, it is likely that the understanding of some counting principles—especially abstraction, one-to-one correspondence, and cardinality principles—favors the perception of collections in terms of discrete quantities. On the one hand, the "discreteness" of collections is outlined by the abstraction and the one-to-one correspondence principles which stipulate that only "distinct entity" property must be abstracted from counted elements and that only one number word can be assigned to each counted object. On the other hand, the cardinality principle implies understanding that the number word associated with the last object counted represents the cardinal of the set. Thus, although the cardinality principle implies understanding that the result of counting represents a quantity, the two others require disregard of perceptual dimensions so that each uttered word corresponds to one entity (one-to-one correspondence) whatever its size, its area, or its identity (abstraction).

To explore the relation between the development of counting skills and the ability to process discrete number properties, we chose to test children at a key age, namely, when they begin to master cardinality principles and to understand the meaning of counting. According to Wynn (1990, 1992b; Fuson, 1992), this conceptual shift occurs around the age of 3. Before this, children are able only to recite the number sequence, without really understanding that the last number word in the counting sequence represents the numerosity of the set. To assess children's counting level, three tasks were used. In the *sequence* task, children were asked to "count as far as possible" (i.e., to recite the number words in the conventional order). In the *how many* task, they were asked to count the objects of a collection and to say how many objects there were. However, this task can be performed by children who adopt a strategy of responding the last number word uttered in the sequence count, without understanding that this word represents the cardinal value of the set. Therefore, in the *give-a-number* task, the child had to give to the experimenter a specified number of objects, which implies understanding of the cardinal meaning of the number words.

The present research

The present study examined the magnitude comparison of preschoolers and addressed two unresolved questions in the literature. First, what is the nature of the processes underlying the performance of infants and preschoolers in numerical tasks? Do young children possess an innate analog mechanism for representing numerosity, do they use a more general and limited mechanism for keeping track of individual entities, or do they rely on correlated perceptual variables? Second, what is the relation between counting abilities and children's performance in quantification tasks?

Recently, Brannon and Van de Walle (2001) attempted to resolve these questions. They investigated preschoolers' quantitative numerical knowledge in a comparison task in which front surface area was controlled. Results showed that children as young as 2 years of age can represent quantitative relations between numerosities if they possess minimal verbal numerical competence. Although this study demonstrates discrete quantitative competence by 2 year olds, it does not conclusively highlight the mechanism underlying such comparison. According to the authors, their results cannot be explained by the operation of an object-file mechanism because the comparison extended to large sets of five and six. However, one term of the comparison was always between 1 and 4, which could have permitted the comparison with such a mechanism. An examination of comparison performance on larger collections would have been useful in answering this question. Second, the authors concluded that underlying the success in their quantitative task was an analog magnitude representation of number. However, the ratio effect predicted by the analog magnitude model was not examined (nor was any distance effect), and the absence of difference in performance between pairs of numerosities that differ in ratio suggests that there is no analog magnitude representation of number. Finally, if children were operating on the basis of such preverbal analog magnitude representation, why did they need minimal counting competence to make the comparison? We argue that numbers begin to be a salient dimension of the environment for children when they start to understand how numerical verbal labels map onto numerosity independent of perceptual variables. Our hypothesis is that minimal counting knowledge is necessary in situations in which numerosity is the only relevant dimension to discriminate or compare sets (i.e., quantifying heterogeneous objects or collections that are very well controlled perceptually), but not in conditions in which perceptual features and number covary. In the latter case, children could rely on perceptual variations and not need to abstract numerosity to resolve the problem. Along with this last hypothesis, we might predict that counting abilities are not a good predictor of the child's success in the conditions with low perceptual control, whereas minimal counting abilities are required to succeed in well-controlled perceptual conditions.

To sum up, the competing theories about the nature of the mechanism underlying numerosity comparison were contrasted according to their different predictions about the size and ratio effects and about the influence of perceptual control. Children's ability to compare numerosity was tested by asking them to choose the larger of two numerical sets that vary in terms of ratio (1/2, 2/3, 3/4) and size (small, medium, and large collections). Thus, we extended the method of Brannon and Van de Walle (2001) by using larger numerosities and by assessing the effect of ratio.[1] Moreover, to assess the influence of perceptual variations on comparison, different perceptual controls were used. In Experiment 1, the different perceptual conditions were intermixed to make number the only relevant cue. In Experiment 2, perceptual conditions were presented separately so that each child completed the task in only one condition of perceptual control. This procedure allowed us to isolate the selective influence of each kind of perceptual control on performance without contamination from the other conditions. The question of the relation between nonverbal numerical abilities and the verbal counting system was explored in both experiments.

[1] Contrary to Brannon and Van de Walle (2001), we chose to use 2-D drawings to better control perceptual variables. Although the choice of 3-D displays would have been closer to the situations children encounter in their environment, we preferred not to add another dimension to control for in our stimuli. Moreover, even when the volume of an object is controlled for, the number of visible sides of that object depends on the child's perspective. So, the simultaneous control of volume and surface of visible sides could have been problematic.

Experiment 1

One procedure used in experiments with infants to ensure that numerosity is the only cue for discrimination has involved introducing a maximum of variations in perceptual features (Starkey et al., 1990; Strauss & Curtis, 1981). This procedure was designed to steer infants' attention to number as a relevant variable. Experiment 1 followed the same idea: pairs of collections with different perceptual controls were presented to the child, who had to choose the array that contained more sticks. Perceptual controls were intermixed randomly to make number the only relevant cue.

Method

Participants

Thirty-three 3-year-old children (mean age 3.5 years, $SD = 2$ months; range 3.1 to 3.9 years; 16 females, 17 males) participated. Three additional children were not included because of an experimenter's error for one child and because the other two failed to understand the task and responded always on the same side (lack of interest). Children were drawn from local preschools in Brussels, Belgium.

Procedure

Children were simultaneously presented with two collections of sticks. They were instructed to select the one that contained more sticks and were informed that the selection of the correct alternative would result in the appearance of a beautiful drawing. This continual reinforcement procedure was used during the whole experiment to avoid the possibility of differential motivation between learning and test phase. The instructions were as follows: "You will see two pictures. You must choose the picture that contains more sticks. If it is the right picture, you will see a beautiful drawing. If it is the wrong picture, there will be nothing."

Children first completed a training phase followed by the test phase and the counting tasks. The experiment was conducted on a PC computer in a quiet room. The experimenter was seated on the right of the child. Each trial followed the same sequence: presentation of the stimulus until response (touching one collection) and immediately after, presentation of a colored drawing (different in each trial) for 2 s in the case of a correct response and presentation of a black screen for 2 s in the case of failure. The presentation of the next trial was ordered by the experimenter after making sure that the child was attentive.

Materials

Comparison materials

Stimuli were pairs of white square cards of 87 mm. The two cards were presented on a black background and were separated by 23 mm. A variable number of black sticks was drawn on each card.

Comparison arrays varied along three ratios (1:2, 2:3, 3:4) and each ratio was represented in three different sizes (small, medium, and large) as displayed in Table 1.

Table 1
Stimulus pairings according to the ratio and the size of the pair

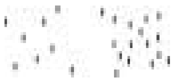
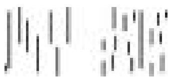
Size	Ratio		
	1:2	2:3	3:4
Small	2–4	2–3	3–4
Medium	4–8	4–6	6–8
Large	8–16	8–12	9–12

Small size corresponded to numerosities that can be apprehended by a mechanism of object-based attention (i.e., up to 4) and large size to numerosities that cannot (greater than 7). Medium-sized pairs were in between and were introduced to balance the stimuli so that in each pair at least one member was presented once as the larger numerosity and once as the smaller throughout the experiment (i.e., 3 vs *4*, *4* vs 8). For that same reason, the pair 12 versus 16 was also presented so that the numerosity 12 was not always the larger of the pair. However, results of this pair were not included in the statistical analysis. For each number pair, the larger numerosity appeared once on the right and once on the left.

To test the hypothesis of a perceptually based comparison, three conditions of perceptual control were used (see Table 2). In the *density* condition, the density of the elements was controlled. In the *contour* condition, the perimeter determined by the most external sticks was controlled (with a 180° rotation between the perimeter of the first array and the perimeter of the second). In these two conditions, the sticks printed on the cards were 2 mm wide and 9 mm high. The third condition, *surface,* controlled the total surface area. One way to equate surface area of the elements of the two arrays is to reduce the size of the elements in the more numerous card. In that case, however, selecting the larger numerosity corresponds to selecting the arrays with the smaller elements. To make this cue less salient, we used sticks of different height. Moreover, in all the stimuli, the smaller and the larger sticks were always the same (smaller stick, 9 mm high; larger stick, 69 mm high). Furthermore, as in the contour condition, the external perimeter of both arrays to be compared was identical. As all sticks were 2 mm wide, the equalization of total surface area also corresponded to an equalization of the sum of the height of the sticks so that differences in the sum of individual contour length were negligible (i.e., contour length ratio varies between 0.95 and 0.99). It should be noted that by controlling the surface area of the sticks, brightness is automatically equated as well. The controlled and uncontrolled perceptual variables for each condition are shown in Table 2.

The test items included 20 stimuli in each perceptual condition (density, contour, and surface). However, as 60 items was thought to be too many for 3-year-olds, the set was divided into three subgroups of 20 stimuli (10 pairs × 2 sides). Each subgroup included the 9 numerosity pairs (3 pairs by perceptual condition) and the filling pair 12–16. The pairs were chosen to include one pair of each ratio (1:2, 2:3, 3:4) and each size (small, medium, large) from each perceptual condition. The three groups of

Table 2
Control of perceptual variables according to the perceptual condition

Perceptual conditions	Perceptual variables				
	Density	External contour length[a]	Total surface area/brightness[b]	Total perimeter[c]	Size of the sticks
Density 	+	–	–	–	+
Contour 	–	+	–	–	+
Surface 	–	+	+	+	–
Heterogeneous size 	–	+	–	–	–

Plus signs indicate that the perceptual variable is controlled. Minus signs indicate that the perceptual variable covaries with number.
[a] Perimeter delimited by the most external sticks.
[b] Sum of the individual item surface.
[c] Sum of the individual item contour length.

stimuli differed by the distribution of the numerosity pairs in the perceptual conditions. Children were randomly assigned to one subgroup of stimuli (i.e., 11 children by subgroup). Table 3 displays the stimuli set of the three subgroups.

Table 3
Distribution of the test pairs in the three groups of children

	Density condition	Contour condition	Surface condition
Group 1	2–4	2–3	3–4
	6–8	4–8	4–6
	8–12	9–12	8–16
Group 2	3–4	2–4	2–3
	4–6	6–8	4–8
	8–16	8–12	9–12
Group 3	2–3	3–4	2–4
	4–8	4–6	6–8
	9–12	8–16	8–12

The stimuli used in the learning phase were different from those of the test phase. They contained large numerosities with a very large ratio of 1:3 so that the difference between the two numerosities was made more salient. Four numerosity pairs were used (7 vs 21, 9 vs 27, 10 vs 30, and 11 vs 33) and were presented six times each (with three presentations of the larger collection on the left). The learning phase was composed of a subset of 24 learning stimuli; that is, 4 numerosity pairs $\times$ 2 sides of correct response $\times$ 3 perceptual conditions (density, contour, surface).

Stimuli were presented in a fixed pseudo-randomized order for both the training trials (with no more than two successive correct responses on the same side) and the test trials (with no more than two successive pairs with the same ratio, the same size, or the same correct response side).

Counting tasks

The three following tasks were used to assess the mastery of the verbal counting system. They were modeled on those used by Wynn (1990, 1992b).

In the *Sequence* task, children were asked to count as far as possible. If required, a prompt (1, 2, ...) was given to the child. When needed, counting was stopped at 16 (the greatest numerosity in our comparison task). The dependent measure was the largest number up to which the child could count correctly or with one error (omission, repetition, or addition of one word), with a maximum of 16.

In the *How many* task, children were invited to enumerate a set of objects. They were presented with sets of 2, 3, 4, or more aligned items (animals printed on 9×2.5-cm cards). The animals were all identical on any given set but differed from one set to the next (dogs, cats, and pigs). The numerosities 2, 3, and 4 were systematically tested in a random order. If the child succeeded, larger numerosities (up to 16) were proposed in a random order until failure. Children were first asked to count sets of items (e.g., "Can you count all the dogs?") and then, immediately after each counting, were asked "so, how many dogs are there?" This allowed us to measure the child's ability to recite the verbal number names, to coordinate the counting and the pointing, and finally, to evaluate his/her understanding that the final word of the counting sequence represents the cardinal of the whole set. Two scores were calculated: one for the counting and one for the response to the "How many" question. The counting score corresponded to the highest numerosity correctly counted (correct enumeration, pointing, and coordination). The score for the "How many" question corresponded to the highest numerosity for which the child could give the cardinal label of the set. The cardinal label was considered correct if it matched the child's counting (even if it was false) and if this cardinal label was not given systematically for all the other numerosities tested.

In the *Give a number* task, children were presented with 10 aligned items (printed on 9×2.5-cm cards) and were asked to give the examiner *n* items (e.g., give me three dogs). Once again, the animals were identical within a trial but different between trials (dogs, cats, and pigs). The numerosities 2, 3, and 4 were proposed systematically in a random order, and larger numerosities (up to 16) were presented for the better performing children, until failure. This task did not require the understanding of quantitative terms such as "many" or "how many," which may be difficult for young children to interpret. This task did not explicitly ask the child to count or to enumerate objects. It was thus possible to see whether the child would spontaneously use counting to solve the task. The score corresponded to the highest numerosity for which the child could give the number of items asked by the examiner. A correct response was credited if the child did not give the same answer to all the other numerosities tested.

Results

Comparison data

Analyses were carried out on accuracy data of the 18 test items. First a repeated-measures ANOVA was run on children's scores with perceptual condition (density, surface, contour) as the within-subject factor. The condition effect was significant, $F(2,\ 64) = 5.67$, $p = .005$, and indicated more accurate performance in the density ($M = 4.15/6$, $SD = 1.30$) and the contour conditions ($M = 4.24/6$, $SD = 1.54$) than in the surface condition ($M = 3.33/6$, $SD = 1.24$; Bonferroni pairwise comparison $p < .05$). Performance exceeded chance in the density, $t(32) = 5.08$, $p < .001$, and the contour conditions, $t(32) = 4.63$, $p < .001$, but not in the surface condition, $t(32) = 1.54$, $p > .05$.

The following analyses were aimed at examining the effects of the ratio and size of the pair in the different situations of comparison. However, as each participant was presented with only one-third of the stimuli, a global ANOVA with all these variables could not be computed. We thus turned to a logit analysis (see Berkson, 1944). This analysis is quite similar to a logistic regression and is particularly well suited to a binary dependent variable. In our case, the dependent variable was the success or failure in one trial. The following four predictors were considered: perceptual condition (−1 for density, 0 for contour, and +1 for surface), ratio (−1 for 1:2, 0 for 2:3, and +1 for 3:4), size of the pair (−1 for small, 0 for medium, and +1 for large), and side of the correct response (+1 for left and −1 for right),[2] plus a random effect for the subject.

We first tested a nearly saturated model that included the four main effects as well as all the first- and second-order interactions between the three main variables: condition, ratio, and size. A total of 594 observations were considered (18 per participant for a total of 33 participants). None of the interactions reached significance but the main effects of condition (estimate parameter −.339: $t(32) = -3.01$, $p = .0051$) and ratio (estimate parameter −.343: $t(32) = -3.04$, $p = .0047$), as well as the intercept (estimate parameter .699, $t(32) = 5.30$, $p < .0001$), were significant.

[2] The side of the correct response was included in the analyses because for Dehaene, Bossini, and Giraux (1998), the number representation is a left–right-oriented mental number line. Our experimental design did not allow us to examine a SNARC effect but we could expect a spatial congruity effect: participant's performance could be more accurate with small numbers on the left and large on the right than the reverse.

To examine more deeply these effects of ratio and condition, an ANOVA was carried out on children's scores with these two variables as within-subject variables, followed by Bonferroni pairwise comparisons. As in the logit analysis, main effects of condition, $F(2,\ 64) = 5.67$, $p = .01$, and ratio, $F(2,\ 64) = 8.06$, $p = .001$, were obtained and no interaction was found, $F(4,\ 128) < 1$. Pairwise comparisons revealed that the surface condition ($M = 3.33/6$) gave rise to lower scores than the density ($M = 4.23/6$, $p < .05$) and the contour conditions ($M = 4.14/6$, $p < .05$) whereas performance was more accurate for the 1:2 ratio ($M = 4.44/6$) than for both the 2:3 ($M = 3.69/6$, $p < .01$) and the 3:4 ratios ($M = 3.6/6$, $p < .005$).

Based on these contrasts, a new way of coding the variables was introduced in a second logit model: The condition was coded +1 for the surface and −1 for both contour and density and the ratio was coded +1 for the 1:2 ratio and −1 for the two other ratios. The model tested here was much simpler than the first one as it included only the main effect of these two new binary variables (interactions, size of the pair, and side of the correct response were not introduced). All the estimated parameters were significant (intercept, .6923, $t(32) = 4.84$, $p < .0001$; condition, −.3376, $t(32) = -3.58$, $p = .0011$; ratio, .3206, $t(32) = 3.21$, $p = .003$) and the fit of this reduced model (−2 log likelihood) was very similar to that of the first nearly saturated model (respectively, fit of 734.9 and 734.8).

In summary, the probability of success for a given comparison pair was a function of the perceptual condition and the ratio: It was lower for the pairs presented in the surface condition than in the contour and density conditions and higher for pairs with a ratio of 1:2 than for those with a ratio of 2:3 or 3:4. The other variables, namely the size of the pair and the side of the correct response, did not affect performance.

Counting data

Table 4 presents the descriptive data for each counting task. Children had limited counting skill. On the Sequence, Counting, How many, and Give a number tasks, 10 (30%), 21 (64%), 19 (58%), and 18 (55%) children, respectively, scored 0. No child could count or give a correct cardinal response for numerosities above 4. Spearman correlations indicated that the Sequence task correlated with the Counting and the Give a number tasks (see Table 4).

Table 4
Children's performance on the counting tasks and Spearman correlations between these tasks and the age of the children in months

	Counting performance		Sequence	Enumerating	How many	Age
	Mean	*SD*				
Sequence	3.91	3.85	1			−.127
Enumerating	1.82	3.22	.378*	1		−.144
How many	0.97	1.19	.160	.127	1	−.121
Give me	1.12	1.32	.480**	.287	0.36	.014

* $p < .05$.
** $p < .01$.

A stepwise regression was conducted on comparison accuracy with the four counting scores and the child's age in months as predictors. Three variables were selected by the regression: the Give a number score was entered first and explained 29.5% of the variance, $F(1,\ 32) = 12.98$, $p = .001$, $MSE = 5.88$; the How many question score increased the explained variance by 14.8%, $F(2,\ 32) = 11.92$, $p < .0009$, $MSE = 4.80$; and finally the counting score added another 7.4% of the explained variance, $F(3,\ 32) = 10.33$, $p < .0009$, $MSE = 4.31$. When separate stepwise regression analyses were conducted for each perceptual condition, the Give a number score was the only variable selected by the analysis for the surface condition and explained 20.5% of the variance, $F(1,\ 32) = 7.97$, $p = .008$, $MSE = 1.27$. For the contour condition, the Give a number score explained 36.1%, $F(1,\ 32) = 17.50$, $p < .0009$, $MSE = 1.57$, of the variance and the Counting score increased the explained variance by 9.8%, $F(2,\ 32) = 12.73$, $p < .0009$, $MSE = 1.37$. For the density condition, the regression was not significant.

In their study, Brannon and Van de Walle (2001) observed that children performed above chance in the comparison task only when they possessed a minimal comprehension of the cardinality concept defined by the capacity to provide at least one correct cardinal response. To examine if such a relation could be replicated here, children were divided into two groups according to their performance in the Give a number task. The 18 children who scored 0 formed the "No cardinal knowledge" group and the remaining 15 children made up the "Some cardinal knowledge" group. These two groups did not differ in age (mean of 3.5 in each group). An ANOVA on comparison score with perceptual condition as within-subject variable and cardinality group as between-subjects factor indicated the known significant effect of perceptual condition, $F(2,\ 62) = 6.22$, $p = .003$, as well as a significant effect of cardinality group, $F(1,\ 31) = 11.72$, $p = .002$. The No cardinal knowledge group ($M = 10.39/18$) performed less accurately than the Some cardinal knowledge group ($M = 13.33/18$). The interaction between cardinality group and perceptual condition was also significant, $F(2,\ 62) = 4.37$, $p < .02$, and indicated that the Some cardinal knowledge group was performing more accurately than the No cardinal knowledge group in the contour and surface conditions especially. Actually, the No cardinal knowledge group performed above chance in the density condition only (density, $M = 4.11/6$, $SD = 1.32$, $t(17) = 3.56$, $p < .005$; contour, $M = 3.44/6$, $SD = 1.46$, $t(17) = 1.29$, $p > .10$; surface, $M = 2.83/6$, $SD = 1.04$, $t(17) = -0.68$, $p > .25$), whereas the Some cardinal knowledge group performed above chance in each of the three perceptual conditions (density, $M = 4.20/6$, $SD = 1.32$, $t(14) = 3.52$, $p < .005$; contour, $M = 5.20/6$, $SD = 1.01$, $t(14) = 8.40$, $p < .0005$; surface, $M = 3.93/6$, $SD = 1.22$, $t(14) = 2.96$, $p < .01$). However, the Some cardinal knowledge group did not exhibit any ratio effect in the surface condition, $F(2,\ 28) = .42$, $p = .66$, despite the fact that their performance was above chance.

In summary, the ability to compare numerosity seems to emerge with some cardinal knowledge. Indeed, only children who possessed some cardinal knowledge performed above chance in the three perceptual conditions. In contrast, children who did not possess such cardinal knowledge performed above chance only in the conditions of low perceptual control.

Discussion

Experiment 1 indicates that mixing perceptual conditions (density, contour, surface) did not lead children to compare sets on a numerical basis. Children failed to discover the number as the only relevant cue because performance was strongly affected by the perceptual control. Children performed above chance when comparing the numerosity of sets for which density or contour was controlled, whereas they answered by guessing when surface was controlled. These results suggest that they used continuous quantitative cues to quantify sets rather than a numerical preverbal mechanism. Furthermore, as comparison performance was more accurate for larger ratios and was unaffected by the size of the numerosity, this perceptual mechanism would be an analog medium for representing amount. Because of their inherently imprecise nature, continuous perceptual properties can be represented only in an approximate way. This explains why performance is subject to Weber's law.

It seems very improbable that children relied on verbal counting to compare the numerosities because no size effect was observed in the comparison task. Furthermore, the Sequence task never predicted the success in the comparison task, and the How many task (Counting and How many question) explained only a very small part of the variance. Yet comparison success was somewhat related to the understanding of the cardinality concept. Indeed, the Give a number performance was the best predictor of the success on the comparison task, especially in the more perceptually controlled conditions (contour and surface). Furthermore, comparison was more accurate in children who possessed a minimal comprehension of the cardinality concept than in those who did not, and this was especially true for the contour and the surface conditions. Indeed, although both groups performed above chance in the density condition, only those who possessed minimal cardinal knowledge did so in the contour and the surface conditions. Nevertheless, individual profiles suggest that minimal cardinality understanding is not necessary to succeed in the contour condition. Indeed, two children were 100% correct in this condition although they had no cardinal concept. The situation is totally different for the surface condition, in which the few children who performed above chance all possessed minimal cardinal knowledge. Having some cardinal knowledge thus seems necessary to succeed in this condition. Nevertheless, having some cardinal knowledge does not seem to be sufficient because six children with some cardinal concept responded randomly in the surface condition.

Experiment 2

Experiment 1 showed that children performed less accurately in the surface condition than in the density and contour conditions. These results indicated that increasing perceptual control led to increasing error rate. The existence of perceptual influence suggests that children do not use numerical representations and rely on perceptual variables to compare collections. However, in this experiment, the stimuli of each perceptual condition were intermixed. This procedure may have favored the use of continuous quantitative cues to compare collections instead of encouraging children to resort to discrete quantification processes. Even if only numerosity consistently predicted the correct response, surface area was not controlled in 2/3 of the trials. The continuous reinforcement procedure may have incited children to use the surface area rather than number to compare sets because surface-based comparison led to a reward in 2/3 of the tests (density and contour conditions).

Therefore, perceptual conditions were used as a between-subjects variable in Experiment 2. The stimuli used were the same as in Experiment 1. Furthermore, a fourth perceptual condition was added to avoid interpretive problems. Indeed, the most difficult condition in Experiment 1 was the surface condition. This could be because the surface area was controlled, but this could also be due to the fact that this was the only condition with heterogeneously sized sticks. To dissociate these two features, the fourth condition included heterogeneous elements in terms of size but without any control of the surface area (see Table 2).

Regarding the density, the contour, and the surface conditions, the predictions are the same as in Experiment 1. Furthermore, the analog numerical models (accumulator and neuronal models) assume a normalization stage in which the perceptual experience is segmented into countable entities to be enumerated regardless of size, color, and so on (see Dehaene & Changeux, 1993, for a description of this processing stage). If this normalization stage is effortful for children, we expect less accurate performance in the two conditions involving heterogeneously sized stimuli (i.e., the surface and the heterogeneous size conditions). Conversely, if children do not process numerosity but rather base their judgment on surface area, they should have difficulty in the only condition which controls for the surface area (i.e., the surface condition). As in Experiment 1, a ratio effect is assumed by the analog numerical model, whereas the object-file model predicts only that comparison is limited to small numerosities. Yet, as surface area perfectly covaries with number in three perceptual control condition (density, contour, and heterogeneous conditions), the amount model also assumes a ratio effect, but only in those three conditions. This would reflect a perceptual ratio effect.

Method

Participants

Sixty 3-year-old children (mean age 3.5 years, $SD = 3$ months; range 2.11 to 3.11 years; 32 females, 28 males) participated in the study. They were drawn from local preschools in Brussels, Belgium. Two additional children who failed to understand the task were not included.

Procedure

As in Experiment 1, each child completed the training phase followed by the test phase and the counting tasks. The procedure was exactly the same as in the first experiment except that children were randomly assigned to one condition of perceptual control (15 children in each condition). There was no age difference between the four subgroups (density, $M = 41$ months, $SD = 3$ months; contour, $M = 43$ months, $SD = 3$ months; surface, $M = 41$ months, $SD = 3$ months; heterogeneous size, $M = 40$ months, $SD = 3$ months; $F(2, 56) = 1.49$, $p > .10$).

Materials

Comparison materials

Stimuli used in the first three perceptual conditions (density, contour, and surface) were the same as in Experiment 1. In the fourth condition, called the *heterogeneous size* condition, each pair was composed of sticks of varying height but with the same smaller and larger sticks (measuring respectively 9 and 69 mm high). Yet, unlike in the surface condition, the total surface area covaried with the numerosity in exactly the same way as it did in the density and contour conditions. For instance, pairs presenting a numerical ratio of 1:2 (2–4, 4–8, or 8–16) also presented a surface ratio of 1:2. Finally, the pairs were also equated in terms of the contour determined by the most external sticks (see Table 2).

The stimuli presented in the learning phase belonged to the same perceptual conditions as those of the test phase. A total of 176 stimuli were thus created (24 learning pairs $\times$ 4 perceptual conditions, and 20 test pairs $\times$ 4 perceptual conditions), which corresponds to 44 items for each perceptual condition. As in Experiment 1, all stimuli were presented in a fixed pseudo-random order for the training (with no more than 2 successive pairs with the same correct response side or from the same perceptual condition) and for the test phase (with no more than 2 successive pairs with the same ratio, the same size, or the same correct response side or from the same perceptual condition).

Counting tasks

After the comparison task, children completed the same counting tasks as in Experiment 1.

Results

Comparison data

Analyses were carried out on the accuracy data of the 18 test items (the pair 12–16 being excluded). Ideally, a preliminary global analysis of variance with perceptual condition as a between-subjects variable and size, ratio, and left–right order for within-subject variables should have been conducted to determine the respective effects of each variable as well as their interactions. However, such an analysis was not made because the number of items included in the experimental design was too small (the choice of including a small number of items in this study was motivated by the limited attention abilities of such young children). Analyses were then conducted according to precise hypotheses.

The first analyses were carried out to determine the kind of mechanism underlying children's comparisons in the different perceptual conditions. If comparisons are based on numerosity, performance should be above chance in all perceptual conditions. An ANOVA was run on children's scores with perceptual condition (density, surface, contour, heterogeneous size) as a between-subjects factor. A main effect of condition emerged, $F(3,\ 56) = 6.37$, $p = .001$, indicating that performance in the surface condition ($M = 10.20/18$) was less accurate than in the three others (density, $M = 14.20/18$, $p = .002$; contour, $M = 14.27/18$, $p = .002$; heterogeneous size, $M = 13.33/18$, $p = .025$). The mean score in each condition was then compared to the score expected by chance (i.e., 9/18) with a one-sample, two-tailed t test. Children in the density, contour, and heterogeneous size conditions performed significantly above chance (density, $M = 14.20$, $SD = 2.51$, $t(14) = 8.02$, $p < .0009$; contour, $M = 14.27$, $SD = 2.60$, $t(14) = 7.83$, $p < .0009$; heterogeneous size, $M = 13.33$, $SD = 3.13$, $t(14) = 5.36$, $p < .0009$), whereas in the surface condition, they responded randomly ($M = 10.20$, $SD = 3.41$, $t(14) = 1.37$, $p > .10$). Thus, the failure in the surface condition is not due to the heterogeneity of the items because performance was above chance level in the heterogeneous size condition. Rather, it appears that increasing perceptual control led to an increasing error rate.

Regarding the ratio and the sizes of the pairs, the analog magnitude model predicts a ratio effect at each numerosity size, whereas the object-file mechanism predicts that only small numerosities could be compared. Finally, if children's comparisons are based on perceptual features rather than on numerosity, a ratio effect is expected in low-controlled perceptual conditions only, in which most perceptual properties, and surface area in particular, covary with number.

These hypotheses were first tested through an item analysis. Stepwise regression analyses were computed in relation to success rate for each item in each condition. The dependent variable was the number of children who correctly answered each item (a maximum of 15 per condition). Four predictors were considered. These included the two independent variables, ratio (1:2, 2:3, 3:4) and size (large = 3, medium = 2, small = 1), as well as two other variables usually taken into account in the adult numerical literature: the numerical distance (the difference between the larger and the smaller numerosities of a pair, ranging between 1 and 8) and the side of the correct response (left = 0, right = 1).

In the density and the heterogeneous size conditions, ratio was the only variable selected by the analysis. It explained respectively 28 and 24% of the variance. In the contour condition, in which two variables were selected, ratio explained 24% of the variance and size increased the explained variance by 18%, $F(2,\ 15) = 5.43$, $p = .02$, $R^2 = .42$. However, this size effect was not confirmed by our subjects' analysis (ANOVA ratio × size).[3] Moreover, this possible size effect did not go in the expected direction and showed more accurate performance for large numerosities ($M = 4.40/6$ for small numerosities, 4.93/6 for medium, and 5/6 for large). Finally, in the surface condition, the regression was not significant. Scores in the surface condition did not

[3] A ratio × size ANOVA revealed a main effect of ratio ($F(2,\ 28) = 6.30$, $p < .01$) reflecting higher scores for 1:2 ratio ($M = 1.78/2$, $SD = .08$) than for 2:3 ratio ($M = 1.53$, $SD = .10$, $p < .05$) and 3:4 ratio ($M = 1.47$, $SD = .10$, $p = .001$) but no effect of size ($F(2,\ 28) = 1.51$, $p = .24$) or interaction ($F(4,\ 56) = 1.13$, $p = .35$) was significant. However, these results must be considered with caution because there are only two items for each combination of size and ratio.

Table 5
Children's performance on the counting tasks and Spearman correlations between these tasks and the age of the children in months

	Counting performance		Sequence	Enumerating	How many	Age
	Mean	*SD*				
Sequence	6.48	4.74	1			.246
Enumerating	4.58	4.11	.630***	1		.154
How many	1.78	3.32	.332**	.332**	1	−.093
Give me	2.35	2.60	.280*	.435***	.362**	.275*

* $p < .05$.
** $p < .01$.
*** $p < .001$.

differ from chance whatever the ratio or the size of the pair, whereas scores in density, contour, and heterogeneous size conditions always did.[4] Thus, it seemed that children's responses in the surface condition were random. We thus decided to discard this condition from further analyses because it only added noise.

The following analysis compared the effect of ratio in the three significant perceptual conditions. A ratio (1:2, 2:3, 3:4) × condition (density, contour, heterogeneous size) ANOVA revealed a main effect of ratio, $F(2,\ 84) = 16.07$, $p < .0009$, but no effect of condition, $F(2,\ 42) = 0.59$, $p = .56$, nor any interaction, $F(4,\ 84) = 0.30$, $p = .879$. Pairwise comparisons (Bonferroni) showed that performance on pairs with a 1:2 ratio ($M = 5.20$) was more accurate than on those with ratios of 2:3 ($M = 4.53$, $p < .0009$) and 3:4 ($M = 4.22$, $p = .002$).

To summarize, ratio was the best predictor of item success in the three conditions that were performed above chance (density, contour, and heterogeneous size). The effect of ratio indicated more accurate performance for pairs with a ratio of 1:2 than for those with a ratio of 2:3 or 3:4. The size variable did not influence children's performance and numerical distance and side of the correct response were not significant predictors. Finally, none of the variables could predict item success in the surface condition, in which performance was completely at random.

Counting data

As in Experiment 1, children's counting skills were limited. Four (7%), 12 (20%), 35 (58%), and 26 (43%) children scored 0 at the Sequence, Counting, How many question, and Give a number tasks, respectively (see Table 5). Spearman correlations indicated

[4] Ratio 1:2: density, $t(14) = 8.81$, $p < .0009$; contour, $t(14) = 10.05$, $p < .0009$; heterogeneous size, $t(14) = 5.14$, $p < .001$; surface, $t(14) = 1.54$, $p = .15$. Ratio 2:3: density, $t(14) = 6.29$, $p < .0009$; contour, $t(14) = 5.53$, $p < .0009$; heterogeneous size, $t(14) = 4.18$, $p = .001$; surface, $t(14) = 2.05$, $p = .06$. Ratio 3:4: density, $t(14) = 4.58$, $p < .0009$; contour, $t(14) = 4.84$, $p < .0009$; heterogeneous size, $t(14) = 3.23$, $p < .01$; surface, $t(14) = 0.81$, $p = .43$. Small size: density, $t(14) = 6.81$, $p < .0009$; contour, $t(14) = 4.84$, $p < .0009$; heterogeneous size, $t(14) = 3.35$, $p = .005$; surface, $t(14) = 1.85$, $p = .09$. Medium size: density, $t(14) = 4.79$, $p < .0009$; contour, $t(14) = 5.85$, $p < .0009$; heterogeneous size, $t(14) = 5.36$, $p < .0009$; surface, $t(14) = -0.19$, $p = .85$. Large size: density, $t(14) = 4.45$, $p = .001$; contour, $t(14) = 6.48$, $p < .0009$; heterogeneous size, $t(14) = 5.28$, $p < .0009$; surface, $t(14) = 1.47$, $p = .16$.

that these tasks were significantly correlated with each other, but did not show clear improvement over the age range considered here except for the Give a number task.

To examine whether children assigned to the different perceptual conditions of the comparison task performed differently on the counting tasks, Kruskal–Wallis non-parametric tests were carried out with perceptual condition as between-subjects factor and scores at the counting tasks as dependent variables. The analyses indicated that children's performance did not differ on the Sequence, $\chi^2(3) = 5.24$, $p = .155$; the How many question, $\chi^2(3) = 1.35$, $p = .721$; and the Give a number tasks, $\chi^2(3) = 3.08$, $p = .380$, but lower scores in Counting were observed for the children in the heterogeneous size condition, $\chi^2(3) = 14.33$, $p = .002$ ($M = 1.97$, 5.07, 6.73, and 4.67, respectively, for the heterogeneous size, the density, the contour, and the surface conditions).

The four scores in the counting tasks and the children's ages in months were used as predictors in a stepwise regression on comparison scores. The Give a number score was the only variable selected by the analysis for the density and the heterogeneous size conditions. It explained respectively 32%, $F(1,\ 14) = 6.07$, $p = .028$, $MSE = 4.64$, and 30%, $F(1,\ 14) = 5.58$, $p = .034$, $MSE = 7.39$, of the variance and this score still correlated with comparison score when age (in months) was controlled (density, $r = .49$, $p = .07$; heterogeneous size, $r = .58$, $p = .03$). In the contour and the surface conditions, the regression was not significant.

As in Experiment 1, two groups were distinguished to examine the relation between comparison performance and the comprehension of cardinality concept defined by the capacity to provide at least one correct cardinal response. The 26 children scoring 0 on the Give a number task formed the No cardinal knowledge group and the remaining 34 children made up the Some cardinal knowledge group. There was a small age difference between the two groups, $t(58) = -2.09$, $p = .041$ ($M = 3.4$ and 3.6, respectively), but the two groups were equally distributed over the four perceptual conditions, $\chi^2(3) = 1.36$, $p = .716$.

A univariate ANOVA was run with perceptual condition and cardinality group as between-subjects factors and comparison score as the dependent variable. In addition to the known effect of perceptual condition, $F(3,\ 52) = 6.16$, $p = .001$, the analysis

Table 6
Mean scores and standard deviations on the comparison task by cardinality group and perceptual condition

Perceptual condition	Cardinality group	*N*	Mean	*SD*
Density	No knowledge	5	12.40	2.88
	Some knowledge	10	15.10	1.85
Contour	No knowledge	6	14.17	2.93
	Some knowledge	9	14.33	2.55
Surface	No knowledge	8	9.25	2.60
	Some knowledge	7	11.29	4.07
Heterogeneous size	No knowledge	7	11.43	3.21
	Some knowledge	8	15.00	2.00

indicated a significant effect of cardinality group, $F(3,\ 52) = 8.48$, $p = .005$, but the interaction was not significant, $F(3,\ 52) = 1.00$, $p = .402$. Lower comparison scores were obtained for the No cardinal knowledge group ($M = 11.58/18$) compared with the Some cardinal knowledge group ($M = 14.09/18$, see Table 6 for the means). However, the performance of the No cardinal knowledge group as well as that of the Some cardinal knowledge group remained at chance in the surface ondition, respectively $t(7) = 0.27$, $p > .25$, and $t(6) = 1.49$, $p > .10$.

Discussion

Experiment 2 replicated the results of Experiment 1. First, a large disparity was observed between the different perceptual conditions: children performed above chance in the density, the contour, and the heterogeneous size conditions but not in the surface condition. Given the success in the new heterogeneous size condition, the failure in the surface condition cannot be explained by the heterogeneity of the sticks' sizes. Thus, we replicated the main results reported in experiments with infants and preschooler children: When surface is not controlled for (density, contour, and heterogeneous size condition), children are sensitive to quantitative changes (Huntley-Fenner & Cannon, 2000; Starkey & Cooper, 1980; Antell & Keating, 1983; Strauss & Curtis, 1981). In contrast, in conditions of strong perceptual control (control of surface area and contour length), children fail to detect numerical changes, as in the Clearfield and Mix (1999, 2001) and Feigenson et al. (2002) studies. Moreover, performance in the surface condition was at chance for each pair size and ratio. These results can be interpreted neither with an analog numerical model nor with an object-file model because in these two cases, children might at least be able to compare some numerosities (respectively the larger ratio or the small collections). We thus hypothesize that children based their comparison on perceptual cues rather than on numerosity per se. More precisely, children would rely on surface area or any related perceptual variables (the sum of the individual perimeters of the elements or the brightness of the array) because the surface condition is the only one which controls for these continuous variables.

Second, in the three conditions in which surface was not controlled, performance was unaffected by the size of the pair, which is not compatible with the predictions based on the hypothesis of an object-file mechanism. By contrast, performance was predicted by the ratio, which supports the hypothesis of an analog mechanism. However, this mechanism would be perceptual rather than numerical because children were unable to process pairs for which perceptual variables were strongly controlled.

Analyses run on counting scores showed that performance on the Sequence and the How many tasks (counting and How many question) never predicted the success in the comparison task. The very low scores that children obtained in these tasks suggest that they did not count the number of stimuli to compare them. Supporting this idea, comparison performance on small numerosities was similar to performance on the large ones despite the fact that many of the children possessed counting abilities only for small numerosities (up to 4).

In contrast, the Give a number performance was a good predictor of the success in the density and heterogeneous size conditions. Furthermore, children who possessed minimal cardinal knowledge performed more accurately in the comparison task than those who did not, regardless of the perceptual condition. Performance of both groups exceeded chance in the density, contour, and heterogeneous size condition, but even children who possessed some cardinal knowledge did not perform above chance in the surface condition. Thus, the cardinality concept is not required in the perceptual conditions in which the comparison can be realized on nonnumerical grounds. But in the surface condition, in which true numerical processing is necessary, minimal cardinal knowledge is not sufficient to succeed in the task.

General discussion

The present study examined the capacity of 3-year-old children to make numerosity comparisons. Two questions are discussed. First, we debate the nature of the mechanism underlying numerosity comparison: do preschoolers possess an innate analog magnitude mechanism, do they use an object-tracking mechanism in quantitative tasks, or do they rely on correlated perceptual variables to compare numerosity? Second, we consider the relation between mastery of the verbal counting system and the development of numerosity representations.

Nature of the mechanism underlying numerosity comparison

The main purpose of this study was to compare the theoretical interpretations of the mechanism underlying magnitude comparison according to their different predictions. The object-file mechanism assumed a size effect, whereas a ratio effect was expected at each numerosity size according to the analog numerical model.

Success in both Experiments 1 and 2 was not influenced by the numerosity size but by the ratio of the pairs. This contradicts the basic assumption of the object-file mechanism, which predicts that only small numerosity could be compared. Conversely, the presence of a ratio effect provides evidence that an analog mechanism is used in the comparison task. However, the nature of this analog mechanism remains to be determined: does it code numerosity or continuous perceptual dimensions?

According to Mix et al. (2002), preschoolers rely on perceptual variables rather than on number to compare collections. In support of this interpretation, in Experiments 1 and 2 we found that the perceptual condition had a strong effect: although children performed at random in the surface condition, they performed above chance in the density, the contour, and the heterogeneous size conditions. As children performed more accurately in the heterogeneous size condition, the lack of homogeneity of the stick size cannot be the only cause of failure in the surface condition. Nor can this failure be due to the procedure used, as the results were reported with a between-subjects design (in Experiment 1) and a within-subject design (in Experiment 2). In this surface condition, performances were random whatever the ratio and the size. Such results cannot be interpreted with an analog numerical model nor with an object-file model but are compatible with the interpretation of Mix et al. Indeed, our results suggest that children in these two experiments based their comparisons on perceptual cues and, more precisely, on surface area (or any confounded perceptual variable such as the sum of the perimeters of the elements or and brightness of the array), rather than on number information.

These results agree with the findings of Clearfield and Mix (1999a, 1999b) and Feigenson et al. (2002) concerning the absence of infants' reactions to number change when continuous variables are controlled. However, these studies tested only the ability of infants to discriminate numerosity of small collections through a habituation paradigm. The present work is noteworthy in that it tested the influence of various perceptual control on the performance of 3-year-olds in a quantitative task demanding the explicit comparison of small and large collections. Our results directly support the hypothesis of Mix et al. (2002) that early quantification processes rely on continuous nonnumerical cues such as area, volume, contour length, or brightness. According to them, these quantification processes are initially nonspecific and undifferentiated, which means that they are used to determine total amount of both discrete and continuous quantities. They also proposed that this amount-based representation would be inherently approximate, which is confirmed by the ratio effect in conditions in Experiments 1 and 2 in which surface area was not controlled. It is still possible that children own an innate numerical representation. However, nothing prevents them from using number information to compare sets in the surface condition and it remains then to be explained why they prefer not to do so.

To summarize, in conditions in which surface was not controlled, performance was above chance, even for the large numerosities. However, when surface was controlled, performance was at chance for each ratio and each size. These results suggest that preschoolers do not base their comparisons on number information but on continuous amount. The ratio effect we found would express the inherent imprecision of the perceptual-based representations. An analog representation of amount thus underlies the children's comparison of visual arrays (in the age range tested here).

One question remains open. In the present study, the size of the collections to be compared was intermixed. At present, much evidence suggests that infants use continuous nonnumerical cues to quantify small sets (Clearfield & Mix, 1999, 2001; Feigenson et al., 2002). However, numerical representation for small numerosity may already be developed in 3-year-olds, but the intermixed presentation of large sets may have inhibited the resort to such representations. In other words, the intermixed procedure may have favored the use of the only process available to quantify both small and large collections, a perceptual one, at the expense of a numerical process, which could have been used only on small sets. Moreover, the exclusive use of large collections during the learning phase may also have encouraged the development of an amount-based strategy. For further investigation, it would be interesting to present small and large collections separately to see if 3-year-olds are influenced by perceptual control when dealing only with small arrays.

Relation between comparison performance and mastery of the verbal counting system

As we failed to highlight the existence of nonverbal numerical representations in most of the preschoolers tested, we wonder how such representations develop and, more especially, if their emergence is related to the acquisition of conventional counting. From this point of view, the 3-year-old age group appears to be appropriate given that children acquire the conventional counting system between 2 and 4 years of age.

We examined whether children could compare numerosities before acquiring even a minimal level of counting proficiency, and whether proficient counters performed more accurately than the others in the comparison task. We predicted that no counting abilities would be needed to succeed in the conditions in which comparison could be based on perceptual variables (density, contour; the situation is less clear for the heterogeneous size condition). In contrast, to succeed in the surface condition, the child must be able to consider each object as a single entity independent of its physical properties (including its extent). As counting acquisition (in particular, the understanding of counting principles) probably contributes to the development of that capacity, we predicted that only children with minimal counting abilities would be able to compare the numerosities in the surface condition.

In Experiments 1 and 2, performance in the Sequence and the How many tasks (enumerating and How many question) was inaccurate and did not predict children's success, regardless of the perceptual condition. Yet comparison performance varied with the child's ability in the Give a number task. Children with minimal cardinal knowledge performed more accurately than those without. If we now consider separately the different perceptual conditions of the comparison task, some interesting results emerge. First, in the conditions in which comparison could be realized on perceptual grounds, children with no cardinal knowledge performed above chance. This was true for the density, the contour, and the heterogeneous size conditions of Experiment 2 and for the density condition of Experiment 1. The situation was less clear in the contour condition of Experiment 1: The No cardinal knowledge group performed at chance level, but some children in this group nevertheless obtained a maximal performance in the comparison task. We thus conclude that minimal cardinality understanding is not needed to succeed in the comparison task in those conditions in which numerosity is confounded with perceptual variables.

With regard to the surface condition in which such a confound was not present, we predicted that counting ability would help the child to develop a sense of "discreteness" and to disregard the physical size of each element. Such knowledge would be useful in dealing with the number comparison task. In Experiment 1, only children in the Some cardinal knowledge group performed above chance level, but in Experiment 2, both children who did and children who did not possess some cardinal knowledge performed at chance level. However, individual patterns indicated that in both experiments children in the No cardinal knowledge group never performed above chance in the surface condition (see Table 7). In other words, the few children who succeeded in the comparison task in the surface condition all possessed a minimal cardinal knowledge (respectively, 6 and 2 children in Experiments 1 and 2).

Table 7
Number of children performing above or at chance level in the surface condition according to their knowledge of the cardinality concept for, respectively, Experiments 1 and 2

Performance in surface condition	Some cardinal concept	No cardinal concept
At chance	9, 5	18, 8
Above chance	6, 2	0, 0

Thus, minimal cardinal knowledge might well be required for real numerical comparison. These results converge with those of Mix et al. (1996; Mix, 1999a, 1999b) as well as Brannon and Van de Walle (2001), who found that only children with some mastery of counting could perform tasks in which true numerical processing is needed. Yet minimal cardinal knowledge does not seem sufficient because more than half the children with some cardinal knowledge still performed at chance level in the surface condition (9 and 5 children, respectively, in Experiments 1 and 2).

As only a few children succeeded in the surface condition, it is difficult to go deeper into the understanding of the relation between counting abilities and numerosity comparison. Many questions have to be clarified. For instance, we should determine which capacity highlighted by the cardinality task is required to compare numerosities. Is it the capacity to consider each element as counting for one entity no matter what its physical properties are? Furthermore, as minimal cardinal knowledge seems to be necessary but not sufficient, other or more capacities might also be necessary and should be defined.

Acknowledgments

This study was supported by the Belgian National Fund for Scientific Research (FNRS), by the Concerted Actions Research of the Belgian French Community (Convention 01/06-267), by Special Research Fund 1999 from the University of Louvain, and by a Research Credit 2001-2002 from the FNRS. We warmly thank Florence Nicol for helping us in the statistical analyses.

References

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

Berkson, J. (1944). Application of the logistic function to bioessay. *Journal of the American Statistical Association, 39*, 357–365.

Bijeljac-Babic, R., Bertoncini, J., & Mehler, J. (1993). How do 4-day-old infants categorize multisyllabic utterances? *Developmental Psychology, 29*, 711–721.

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition, 83*, 223–240.

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

Canfield, R. L., & Smith, E. G. (1996). Number-based expectations and sequential enumeration by 5-month-old infants. *Developmental Psychology, 20*, 493–523.

Clearfield, M. W., & Mix, K. S. (1999). Number versus 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.

Cooper, R. G. (1984). Early number development: Discovering number space with addition and subtraction. In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 157–199). Hillsdale, NJ: Erlbaum.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*, 371–396.

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

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

Fuson, K. S. (1992). Relations entre comptage et cardinalité chez les enfants de 2 à 8 ans. 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.

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

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175–219.

Koechlin, E., Dehaene, S., & Mehler, J. (1997). Numerical transformations in five-month-old human infants. *Mathematical Cognition, 3*, 89–104.

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

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9*, 320–334.

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

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

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

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

Moore, D., Benenson, J., Reznick, J. S., Peterson, M., & Kagan, J. (1987). Effect of auditory numerical information on infants' looking behavior: Contradictory evidence. *Developmental Psychology, 23*, 665–670.

Simon, T. J. (1997). Reconceptualizing the origins of number knowledge: A non-numerical account. *Cognitive Development, 12*, 349–372.

Simon, T. J., Hespos, S. J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development, 10*, 253–269.

Starkey, P., & Cooper, R. G. (1980). Perception of numbers by human infants. *Science, 210*, 1033–1035.

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.

Strauss, M. S., & Curtis, L. E. (1984). Development of numerical concepts in infancy. In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 131–155). Hillsdale, NJ: Erlbaum.

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.

Uller, C., Carey, S., Huntley-Fenner, G., & Klatt, L. (1999). What representations might underlie infant numerical knowledge? *Cognitive Development, 14*, 1–36.

van Loosbroek, E., & Smitsman, A. W. (1990). Visual perception of numerosity in infancy. *Developmental Psychology, 26*, 916–922.
Wynn, K. (1990). Children's understanding of counting. *Cognition, 36*, 155–193.
Wynn, K. (1992a). Addition and subtraction by human infants. *Nature, 358*, 749–750.
Wynn, K. (1992b). Children's acquisition of the number words and the counting system. *Cognitive Psychology, 24*, 220–251.
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.