



One, two, three, many – Subitizing in active touch

Myrthe A. Plaisier*, Wouter M. Bergmann Tiest, Astrid M.L. Kappers

Helmholtz Institute, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

ARTICLE INFO

Article history:

Received 9 October 2008

Received in revised form 21 April 2009

Accepted 22 April 2009

PsycINFO classification:

2320

Keywords:

Active touch

Subitizing

Numerosity judgement

ABSTRACT

'Subitizing' refers to rapid and accurate judgement of small numbers of items, while response times and error rates increase rapidly for larger set-sizes. Most enumeration studies have been done in vision. Enumeration studies in touch have mostly involved 'passive touch', i.e. touch without active exploration. In daily life a much more common situation is that of 'active touch', e.g. when we count the number of coins in our pocket. To investigate numerosity judgement in active touch, we let subjects haptically explore varying numbers of spheres. Our results show that enumeration for up to 3 items is more efficient than for larger numbers of items. We also show that enumeration in this regime was not performed through estimation. Furthermore, it is shown that numerosity information was accessed directly and not through mass or volume cues. Not only do our results show that a haptic version of subitizing exists in active touch, they also suggest similar underlying enumeration mechanisms across different modalities.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

From visual studies it is known that people judge the numbers of items on a display rapidly, accurately and almost effortlessly up to a certain number. This phenomenon is known as subitizing (Kaufman, Lord, Reese, & Volkmann, 1949). With increasing set-sizes, enumeration becomes error prone and response times (RT) increase markedly. Consequently, subitizing is characterised by a sharp upward bend in the slope of the RTs and error rates as a function of set-size. The location of the bend depends on the stimulus, but in vision it is generally at about four items (e.g. Atkinson, Campbell, & Francis, 1976; Mandler & Shebo, 1982). A total of three processes for numerosity judgement can be distinguished. First, for small numerosities there is the efficient and accurate process labelled 'subitizing' as described above. Secondly, for larger numerosities a more time-consuming and error-prone process is used which is referred to as 'counting'. Finally, there is the efficient process of 'estimation' for approximate numerosity judgement. The term 'enumeration' refers to numerosity judgement in general through any of these processes.

Outside the visual modality, numerosity judgement experiments in audition have also reported evidence for subitizing (Camos & Tillmann, 2008; Ten Hoopen & Vos, 1979). In this case items are usually presented sequentially and not simultaneously as is often done in vision. There has been much debate on whether two separate mechanisms are involved for subitizing and counting.

Recently several brain imaging studies have focussed on this question. Pasini and Tessari (2001) suggested left hemispherical specialisation for subitizing and right hemispherical specialisation for counting. A study by Piazza, Mechelli, Price, and Butterworth (2006) reported left hemispherical specialisation for approximate numerosity judgement in both vision and audition. If a similar or even a single mechanism underlies both visual and auditory numerosity judgement, it is likely that this mechanism also extends to the haptic modality.

In touch, enumeration studies have been mostly restricted to 'passive touch', i.e. touch without active exploration. One study in which subjects had to judge the number of fingers stimulated with pins reported subitizing (Riggs et al., 2006), but no subitizing was found in a study where subjects had to report how many vibrators were distributed over the body surface (Gallace, Tan, & Spence, 2006). This raises the question whether subitizing only occurs in touch when stimuli are presented to the separate fingers. In a follow-up study where vibrators were presented to the subjects' fingers, Gallace, Tan, and Spence (2008) again reported no indication for subitizing. However, for both presentation to the separate fingers and presentation distributed over the body surface, error rates were extremely high (up to 90%). These results show that stimulation of the separate fingers does not necessarily lead to subitizing, but it is also possible that vibrators are not a suitable stimulus for investigating tactile numerosity judgement. It is still not unlikely that stimulation of the separate fingers represents a special case in haptic numerosity judgement for which subitizing can occur. Recent studies have shown that there are interactions between spatial and number representations in the parietal cortex

* Corresponding author. Tel.: +31 30 253 2807; fax: +31 30 252 2664.
E-mail address: M.A.Plaisier@uu.nl (M.A. Plaisier).

(see Hubbard, Piazza, Pinel, & Dehaene (2005) for an overview). It has been shown that areas involved in number processing partially overlap with those involved in finger movements (Pesenti, Thioux, Seron, & De Volder, 2000). Furthermore, finger motor circuits are activated during cognitive tasks such as enumeration and can facilitate cognitive processing (Andres, Seron, & Olivier, 2007; Carlson, Avraamides, Cary, & Strasberg, 2007). Since motor circuits are activated during number processing, an even more important question than that of the importance of stimulating separate fingers is whether subitizing also occurs in active touch.

In daily life, we usually explore objects through active touch. Allowing to actively explore enables the subjects to adopt the most efficient exploration strategy (Lederman & Klatzky, 1987). When exploratory movements are restricted, haptic information processing can be impaired (e.g. Lederman & Klatzky, 2004). Therefore, experiments using active or passive touch do not necessarily yield similar results. We investigated whether subitizing occurs in active touch by letting subjects enumerate varying numbers of spheres grasped in the hand. Response time and error rates were then recorded as a function of the numbers of items. First a numerosity judgement experiment was carried out to investigate whether two regimes exist in this type of numerosity judgement task. A second experiment was performed to investigate the role of relative discriminability between the presented numbers of items. In the last experiment the role of volume and mass cues was investigated. Finally, we introduce a model to predict response times for numerosity estimation and determine whether this model can describe the data from the second and third experiments.

2. General methods

2.1. Subjects

Thirty paid participants (13 male, 22 ± 3 years) were randomly distributed over three experimental groups. One participant was left-handed and all others were right-handed according to Coren's test (Coren, 1993) and none of them had any known hand deficits. All subjects were naive as to the purpose of the experiment and they were treated in accordance with the local guidelines.

2.2. Stimuli and set-up

The stimuli consisted of brass spheres, which were suspended from wires (Fig. 1a). The wires were flexible enough for subjects to freely lift and rearrange the spheres in the hand. In Experiments 1 and 2 the numbers of spheres was varied and each sphere had a diameter of 1.86 cm and a mass of 29 g. The size of these spheres

was chosen such that varying number of spheres could be held comfortably in one hand, while they were large enough to easily resolve the individual spheres. In Experiment 3 sphere size was varied and the set included spheres with diameters of 1.86 cm, 2.34 cm, 2.95 cm, 3.72 cm and 4.69 cm. Response times were measured using a custom built device. Time measurement was started automatically when a subject touched the stimuli and it was terminated through a vocal response. The RTs were measured with an accuracy of 10 ms. For technical details about this device, see Plaisier, Bergmann Tiest, and Kappers (2008).

2.3. Design

Subjects were blindfolded and wore earplugs to eliminate sound cues. They were instructed to grasp the stimuli with their dominant hand and respond by calling out the correct number of spheres (Experiments 1 and 2) or sphere size (Experiment 3) as fast as possible. There were no restrictions on exploration strategy nor on hand movements, other than having to initially grasp all items simultaneously. After grasping all items they were allowed to release spheres from their hand during a trial. Whenever an incorrect response was given, the experimenter informed the subjects of what the response should have been and the response time was discarded. Before the experiment was started, a minimum of 20 practice trials was performed until 10 in a row were correct. The number of practice trials never exceeded 30 trials. Each subject performed 25 trials per number of spheres or sphere size. There was a 5-min break after 50 min and none of the experiments took more than 90 min.

2.4. Analysis

Strategy shifts in cognitive tasks can be detected by regression of a model consisting of multiple linear parts with unknown change points to, for instance, response times (Luwel, Beem, Onghena, & Verschaffel, 2001). To determine the location of a possible transition point in our data, a bilinear function for the response time T as a function of the number of items N consisting of two linear regimes with a discrete transition was used:

$$T(N) = (r_1 N + c_1)H\left(\frac{c_2 - c_1}{r_1 - r_2} - N\right) + (r_2 N + c_2)H\left(N - \frac{c_2 - c_1}{r_1 - r_2}\right) \quad (1)$$

here $H(N)$ is the Heaviside step function and r_1 and r_2 are the slopes, while c_1 and c_2 represent constant offsets. Regression of this function was performed on the response times averaged over subjects from Experiment 1. We also checked whether a bilinear function

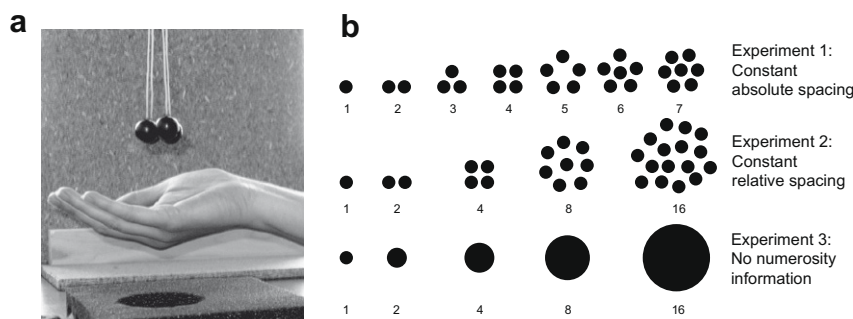


Fig. 1. Experimental design. (a) Blindfolded subjects placed their dominant hand in a holder below the stimulus and grasped upwards. Time measurement started automatically when a subject touched the stimulus. Subjects had to call out the number of spheres, or the number of equivalent small spheres, as fast as possible, terminating the measurement. (b) In Experiment 1 the number of spheres increased linearly from 1 to 7, while in Experiment 2 the number of spheres doubled repeatedly from 1 to 16. In Experiment 3 subjects were presented with single spheres having masses and volumes equivalent to the 1–16 spheres in the previous experiment.

described the data better than two other models: a bilinear function with slope $r_1 = 0$ and a linear function.

Note that a linear function has 2 free parameters, while the bilinear function has 4 free parameters and the bilinear function with the first slope set to zero has 3 free parameters. Regression of a function with more free parameters is more likely to yield a larger R^2 value, so we cannot compare the performance of these models by looking at the R^2 values only. To determine how well the three functions described the data while taking into account the differences in the numbers of free parameters, the Akaike information criterion (AIC) was used (Akaike, 1974). With this method the best function can be selected by looking at how much information is lost when a certain function is used to describe the data. The model for which this information loss is the smallest is then selected as the best model. This calculation is based on the sums of squares (SS), the number of data points (n) and the number of free parameters (k). Here we used AIC with small sample-size correction (AIC_c), which is defined as:

$$\text{AIC}_c = n \ln \left(\frac{\text{SS}}{n} \right) + 2k + \frac{2k(k+1)}{n-k-1} \quad (2)$$

From these AIC_c values Akaike weights (w_a) can be calculated. These w_a represent the relative probability that a certain function in a set of M functions describes the data best. The function with the w_a closest to unity is then determined to perform best. These values are obtained by the following equation:

$$w_a^i = \frac{e^{-\frac{1}{2}\text{AIC}_c^i}}{\sum_{m=1}^M e^{-\frac{1}{2}\text{AIC}_c^m}} \quad (3)$$

3. Experiment 1: constant absolute spacing

This experiment was a haptic version of a 'classic' visual numerosity judgement experiment in which subjects are presented with varying numbers of items and have to enumerate them. If haptic numerosity judgement is similar to that in vision we expect to find a sharp upward bend in both response times and error rates.

3.1. Methods

Subjects were presented with 1, 2, 3, 4, 5, 6 or 7 items (Fig. 1b). Numerosities larger than 7 were not presented, because all subjects had to be able to hold all the spheres comfortably in one hand for the whole numerosity range. They also had to move the spheres in the hand without the spheres falling out of the hand unintentionally. Instructions were as described in the General Methods.

3.2. Results

The averaged RTs as well as the error rates are shown in Fig. 2. Note that no incorrect responses were given up to 3 items and that the error rates increased for larger numbers of items. If errors were made the presented numbers of spheres was always confused only by numerosities of one sphere less or one sphere more. The low error rates indicate that subjects were able to perform the task very well and they did not systematically under- or over-estimate the number of spheres. The response times seem to show two regimes. To test this, regression of a linear function, a bilinear function and a bilinear function with the first slope set to zero was performed on the averaged response times. Data points were weighted according to their standard error. Regression yielded $R^2 = 0.78$ for the linear model, $R^2 = 0.99$ for the bilinear model and $R^2 = 0.99$ for the bilinear model with the first slope set to zero. As pointed out before, we

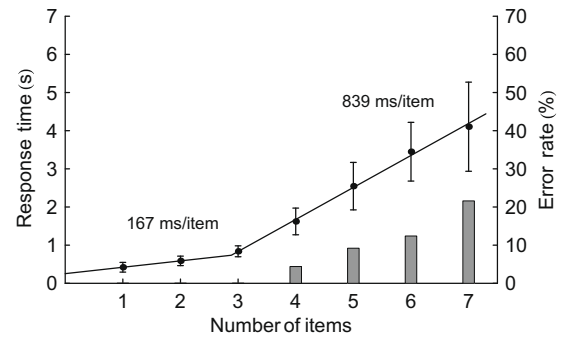


Fig. 2. Experiment 1: Constant absolute spacing. Response times (dots \pm s.e.m.) and error rates (grey bars) as a function of the number of items averaged over subjects. Note that no errors were made below 4 items. The solid line represents regression of the bilinear function and the slope values are indicated in the figure.

cannot conclude from these R^2 values that the bilinear functions perform best, because the functions differ in the number of free parameters. To take the different numbers of free parameters into account, Akaike weights were calculated. This yielded $w_a^1 = 0.97$ for the bilinear function, $w_a^2 = 0.03$ for the bilinear function with the first slope set to zero and $w_a^3 = 0.0003$ for the linear function. This shows that the bilinear function describes the data best with a probability close to one.

Regression of the bilinear function to the response times averaged over subjects yielded a slope transition at 2.9 ± 0.3 (s.e.m.) items. The reported standard error is the value that follows from the regression procedure. The slope of the first linear regime was determined at 167 ms/item, while for the second regime the slope was determined at 839 ms/item. To check how representative the values from the fit to the average data are, the single subjects' data were also analysed by regression of the bilinear function. The location of the transition point, for instance, may vary among subjects. From the fits to the single subjects data it followed that on average the transition point was located at 3.4 ± 0.8 items. Here, the reported confidence interval is the standard error of the values found for the different subjects. The slope values for r_1 and r_2 were 170 ± 30 ms/item and 976 ± 70 ms/item, respectively. The fitted slopes of the first linear part were significantly smaller than those for the second part (paired-samples t -test on the slopes of the individual subjects' RTs, $t_9 = 13$, $P < 0.001$). In Fig. 2 it can be seen that only two data points lie in the first regime. Note that we did not force the transition point from the first regime to the second in between 2 and 3 items. Instead, we minimised the total sum of squares of the model as a whole and not the two linear parts separately. Therefore, the first slope is not simply a connection of the first two data points but the result of a fitting procedure. Furthermore, regression of this model to the single subjects' data, showed that the location of transition point varied between subjects and was often in between 3 and 4 items. The average value of the slope in the first regime for the single subject fits is almost the same as the slope that resulted from the fit to the response times averaged over subjects. This shows that the value of the first slope found through this procedure is robust.

In some studies the largest number of items (in this case 7 items) is not taken into the analysis because of possible endpoint effects. It is common in numerosity judgement data that for the highest numerosity, response times are lower than would be expected from the trend in the previous data points (e.g. Trick & Pylyshyn, 1993; Watson, Maylor, & Bruce, 2007). Although we did not explicitly inform the subjects of what was the largest numerosity, it is likely they noticed it during the experiment. However, the response time for 7 items did not significantly deviate from the trend

in the previous three points¹ and the regression procedure yielded similar parameters if the response time for 7 items was excluded.

3.3. Discussion

Our results show that the response times as a function of the number of items show a sharp upward bend at about 3 items. This relation is better described using a bilinear function than a single linear function, or a bilinear model with the first slope set to zero. This shows that the subitizing slope is not zero. This is in agreement with visual numerosity judgement experiments, in which the subitizing slope is also larger than zero (e.g. Akin & Chase, 1978; Oyama, Kikuchi, & Ichihara, 1981; Trick, 2008; Trick & Pylyshyn, 1993).

Analysis of the single subject data showed a transition point on average between 3 and 4 items and the slopes for the first part of the bilinear function were significantly smaller than those of the second part. Together with the error rates, this provides strong evidence that there are two regimes when making numerosity judgements using active touch. Since there was an upward bend in the response times and also error rates increased after the bending point, it can be assumed that subjects were *counting* for numerosities larger than 3 items. For small numerosities (<4) subjects used a more efficient and accurate enumeration strategy. Note that we chose the sphere size such that up to 7 could be comfortably held in the hand so the spheres not fitting in the hand was no reason for a change in performance after 3 spheres. A possible explanation is that subjects used an estimation strategy to judge small numerosities very efficiently. Note that while the absolute differences between subsequent numerosities were constant, the *relative* differences in numerosity, mass and volume decreased with the number of items, making discrimination progressively harder. A second experiment was performed to investigate the role of relative discriminability on numerosity judgement.

4. Experiment 2: constant relative spacing

In vision, discriminability differences have been suggested as an explanation for the existence of two regimes (Van Oeffelen & Vos, 1982). Furthermore, Ross (2003) suggested that subitizing might be explained from a Weber fraction for visual discrimination of numerosities. When the relative difference between two subsequent numerosities is smaller than the Weber fraction, subjects might switch from one enumeration strategy to another.

To investigate the influence of relative discriminabilities between the presented numerosities, relative differences between subsequent numerosities were kept constant in this experiment. Each numerosity differed by a factor of two from the previous numerosity in the range. A factor of two was the largest difference between two subsequent numerosities that were presented in Experiment 1. If indeed the transition from efficient enumeration to counting that was found in Experiment 1 was caused by the decreasing relative differences between the presented numerosities, we hypothesise that subjects will use the same enumeration strategy over the whole numerosity range in this experiment. Performance should then be roughly constant over the whole numerosity range in the present experiment.

4.1. Methods

Stimuli consisted of 1, 2, 4, 8 or 16 spheres (Fig. 1b), so that the relative volume and mass differences between the subsequent

numbers of spheres were constant. In this case it was no problem to present more than 7 items, in contrast to Experiment 1, because pilot experiments showed that subjects used the same exploratory strategy for the whole numerosity range. They only grasped the spheres and did not need to make exploratory movements. Before the experiment started subjects were informed of which numbers of items could be presented.

4.2. Results

Fig. 3a shows that the response times for up to 16 items were all below 1 s and error rates were low for all numerosities. This indicates that enumeration was facilitated by increasing the relative mass and volume differences. The effect of numerosity was significant (repeated measures ANOVA, $F_{4,36} = 58$, $P < 0.001$) and Bonferroni-corrected pair-wise comparisons showed that enumeration was significantly faster for up to 4 items than for 8 and 16 items. So, RTs were still significantly lower in the first part of the stimulus range. For comparison, Fig. 3b shows the results from this experiment together with the response times from Experiment 1 (in grey). It can be seen that for 1 and 2 items response times are comparable between the experiments, but for larger numbers of items the response times from experiment 1 increase rapidly.

4.3. Discussion

The results clearly show that when the numbers of items were spaced further apart, enumeration of large numbers of items (8 and 16) was facilitated markedly. For numerosities larger than 2 the response times from Experiment 1 increased much more rapidly than in the present experiment. This indicates that an estimation process was used to judge these numerosities and subjects did not count the individual items. However, RTs were not constant over the different numerosities. The significantly lower response times for small numbers of items (≤ 4) suggest that a still more effi-

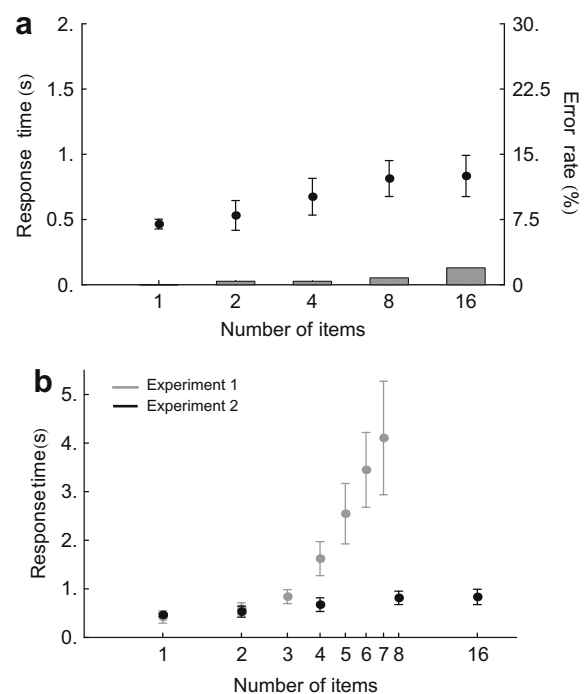


Fig. 3. Experiment 2: Constant relative spacing. (a) Response times (dots \pm s.e.m.) and error rates (grey bars) as a function of the number of items averaged over subjects. (b) The response times from Experiment 2 plotted together with the response times from Experiment 1 (in grey).

¹ The response time for 7 items deviated 0.2 s.e.m. from a linear fit through the data for 4, 5 and 6 items.

cient enumeration process was applied for small numbers of items. This result suggests that subjects did not use the same enumeration process over the whole numerosity range contradicting the hypothesis that smaller relative differences between the numbers of items at the low end of the stimulus range is the cause of the two regimes found in Experiment 1. Possibly, subjects used volume and mass cues to estimate numerosity for the large number of items, but used a different, more efficient, strategy for small numbers of items in the present experiment. This strategy should then not be based on mass and volume cues, but on numerosity information. Experiment 3 was carried out to investigate the role of mass and volume estimation.

5. Experiment 3: no numerosity information

To investigate the role of mass and volume cues, subjects were deprived of numerosity information and only mass and volume cues were available. If the efficient performance for small numerosities, or rather small masses and volumes, remains, then it would be an effect of mass and volume estimation. If, on the other hand, the efficient performance disappears, numerosity information would be accessed directly and not through volume and weight estimation.

5.1. Methods

In this experiment stimuli consisted of single spheres having volumes and weights equivalent to the varying numbers of spheres in the progressive-spacing experiment. The spheres were labelled 1, 2, 4, 8 and 16 equivalent to the numbers of spheres in the previous experiment (see Fig. 1b) and subjects were instructed to respond with the correct label. With the removal of numerosity information, this task has in fact become a categorisation task.

5.2. Results

The results for this experiment are shown in Fig. 4. RTs for all numbers of items except 16 were significantly higher than in Experiment 2 ($t_{18} \geq 2.4, P \leq 0.034$) and particularly for 1 and 2 items did the error rates increase substantially. In error trials the presented sphere was always confused with the previous or the next sphere size in the range (e.g. sphere 8 was confused with either sphere 4 or 16 and sphere size 2 was confused with either 1 or 4). The effect of sphere size was significant (repeated measures ANOVA, $F(4, 36) = 7.9, P < 0.001$). This can be seen from Fig. 4, in which the response times at both ends of the stimulus range are smaller.

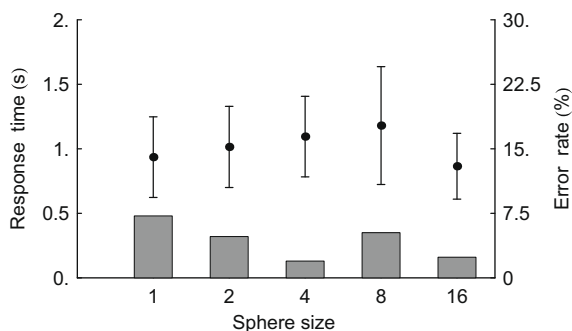


Fig. 4. Experiment 3: No numerosity information. Response times (dots ± s.e.m.) and error rates (grey bars) as a function of sphere size averaged over subjects.

5.3. Discussion

The results show that in the absence of numerosity information, performance in the first part of the stimulus range deteriorated and the highest error rate was found for the sphere size which was equivalent to 1 item in Experiment 2. The more efficient performance for small numerosities found in Experiment 2 disappeared. Although there was no special regime found in the present experiment, performance was not completely constant over the whole stimulus range. A model based on relative discriminability was designed to describe the expected pattern in the response times for numerosity judgement through an estimation process.

6. Estimation model

If subjects used an estimation strategy, we would expect discriminability between the different stimuli in the set to play a role. Discriminability can be assumed to be proportional to the perceived difference between two stimuli. Perceived difference in magnitude usually obeys Fechner's law. According to Fechner's law the perceived difference between two quantities scales with the logarithm of the ratio of these quantities. Therefore, it can be assumed that the discriminability d between quantities x_1 and x_2 is proportional to:

$$d(x_1, x_2) \propto \left| \log \frac{x_1}{x_2} \right| \quad (4)$$

In Experiments 2 and 3, the subjects knew which numerosities could be presented. If they used an estimation strategy, it is expected that when judging a stimulus they considered each of the possibilities weighted according to their discriminability. The total response time is then expected to be inversely related to the sum of the discriminabilities. This is then given by:

$$RT(N) = a + \frac{b}{\sum_{n=i}^j \left| \log \frac{N}{n} \right|} \quad (5)$$

where N is the numerosity that is presented, n is an iterator which runs from the smallest numerosity in the set (i) to the largest one (j) over all numerosities in the set. Free parameters a and b scale the offset and width of the function.

This model was fitted to the response times found in both Experiment 2 and Experiment 3 using nonlinear regression. Each data point was weighted according to the corresponding standard error. Fig. 5 shows the best fit for both sets of response times. The data from Experiment 3 are described fairly well by this model ($R^2 = 0.77$). In this case the model also performs much better than a straight line ($R^2 < 0.01$). On the other hand, the model is clearly not suitable to describe the data from Experiment 2 ($R^2 = 0.34$). In fact, it performs worse than a straight line ($R^2 = 0.86$). Note that in this case we can directly compare R^2 values between a straight line and our model, because both have the same number of free parameters.

6.1. Discussion

The estimation model based on discriminability describes the pattern in the response times when there was no numerosity information fairly well. Note also that this model predicts end effects, i.e. response times at the ends of the stimulus range are smaller. This is often found in numerosity judgement experiments. However, this model cannot describe the data from Experiment 2 in which numerosity information was present. Therefore, the conclusion has to be that other processes than just mass and volume

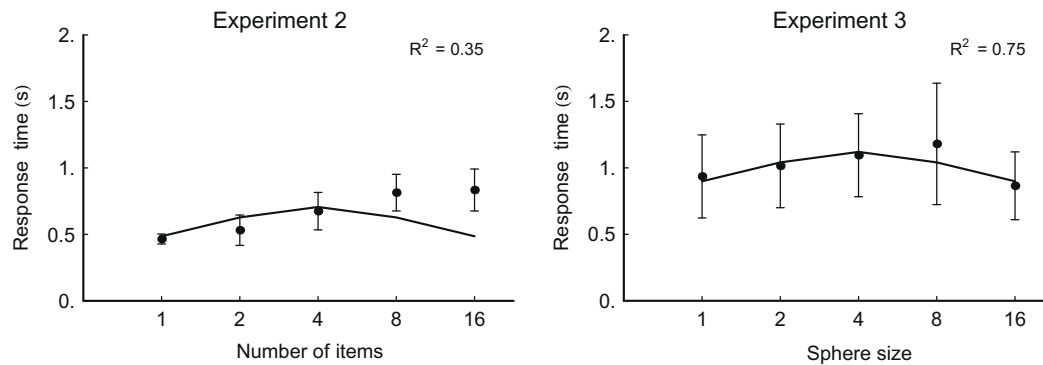


Fig. 5. Best fit of the estimation model to the response times averaged over subjects for both Experiment 2 with numerosity information present and Experiment 3 without numerosity information. R^2 values are indicated in the figure. It is clear that the estimation model does not fit the data from Experiment 2. This indicates that a different enumeration mechanism is used when numerosity information is present.

estimation play a role in this experiment.² Response times for the largest numerosities in Experiment 2 were similar to those for the largest spheres in Experiment 3. Also, there seems to be an endpoint effect for the largest numerosity in Experiment 2 similar to the endpoint effect for the largest sphere in Experiment 3. This suggests that for the largest numerosities (8 and 16) in Experiment 2 subjects used an estimation process like in Experiment 3. For the smallest numerosity in Experiment 2 there is no endpoint effect for the smallest sphere in Experiment 3 indicating that indeed in this regime a different enumeration process was used. The estimation model assumes the same estimation process for all stimuli in the set. Since this model cannot be used to describe the pattern in the response times from Experiment 2, this suggests that subjects used different enumeration processes for small and large numerosities in that experiment. Note that in line with this interpretation, the estimation model cannot describe the data from Experiment 1 either. Also in this case different mechanisms are used over different parts of the range of numerosities as there is a transition from subitizing to counting.

7. General discussion

In the visual domain there is still debate on what type of process subitizing actually is. Several explanations have been proposed for this efficient enumeration of small numbers of items. Mandler and Shebo (1982) showed that presenting dots in canonical patterns can increase the subitizing range from 4 to 6 items. Pattern-based explanations of subitizing suggest that for small numbers of items pattern recognition plays a role. Similarity between different numbers of items is sometimes linked to this pattern-based explanation. Logan and Zbrodoff (2003) showed that when subjects had to rate the similarity between different numbers of items, they rated small numbers of items very dissimilar to each other whereas larger numbers of items tended to be rated increasingly similar. Visual numerosity discrimination experiments have shown that there is a constant Weber fraction of 25% for number discrimination (Ross, 2003). This could explain why there is a transition from subitizing to counting at about 4 items, because this is where the relative difference between two subsequent numerosities drops below the 25% discrimination threshold. Recently it has been shown that this explanation does not hold (Revkin, Piazza, Izard, Cohen, & Dehaene, 2008). In the present study we found a more efficient enumeration regime

for small numerosities when the relative differences between subsequent numerosities were always a factor of two (Experiment 2). This indicates that also in haptic numerosity judgement, the efficient enumeration of small numbers of items is not caused by large relative differences between the presented numerosities. Such a regime of efficient performance was not present when numerosity information was absent (Experiment 3). In that case a model for predicting response times based on Fechner's law described the response times very well, while this model could not describe the response times from Experiment 2. This shows that numerosity information was accessed directly for small numerosities and not through mass or volume estimation.

Subitizing is sometimes assumed to be a purely parallel process yielding response times that are independent of the number of items. This is, however, not generally found in visual studies. In visual numerosity judgement studies where subjects were not forced to answer within a certain time interval, the slope of the response times as a function of the number of items in the subitizing regime is generally 40–100 ms/item, while the counting slope is between 250 and 350 ms/item (e.g. Akin & Chase, 1978; Oyama et al., 1981; Trick & Pylyshyn, 1993; Trick, 2008). In visual search studies, a continuous range of response time slopes is found from purely parallel to purely serial search (Wolfe, 1998). This is generally explained using search models that assume that visual searches are performed through a combination of parallel and serial processes (e.g. Cave & Wolfe, 1990; Duncan & Humphreys, 1989; Theeuwes, 1993). If we extrapolate this idea to numerosity judgement, then the subitizing slope is expected near the parallel end of the range and the counting slope on the serial end of the range of slopes. In a previous study we have measured a range of search slopes in the haptic modality using the same set-up as in the present study (Plaisier, Bergmann Tiest, & Kappers, 2009). Like in visual search, also a large range of slope values was found. These ranged from highly efficient search for a tetrahedron among spheres (88 ms/item, target absent), to very inefficient for an ellipsoid among spheres (1200 ms/item, target absent). Because in numerosity judgement always all items have to be processed, comparison to target absent search slopes is appropriate here. The subitizing slope we found in Experiment 1 is somewhat larger than the visual subitizing slopes (167 ms/item), but it is near the parallel end of the range of haptic search slopes. The counting slope found in Experiment 1 (839 ms/item) is on the serial end of the range of search slopes. So, although the haptic subitizing and counting slopes found in the present study may be larger than generally found in vision, they are in agreement with the values one would expect from the range of slopes found in haptic search. Note also that our subitizing slope is substantially smaller than the tactile subitizing slope reported by Riggs et al. (2006) (270 ms/item).

² Note that weighted regression of the model was used, which means the response times were weighted according to their standard error. Since the standard error for small numerosities was smaller than for the larger numerosities, the distance between the response times and the fit for Experiment 2 are smallest for the first part of the stimulus range. However, this does not mean that the model can describe the data for small numerosities, but not for the larger numerosities.

The fact that the slope for small numerosities is larger for the haptic modality than in vision does not rule out that similar mechanisms underlie numerosity judgements in this regime, especially since in search tasks response time slopes are generally larger in the haptic modality than in vision.

Information extraction is quite different between these two modalities and before items can be enumerated through any mechanism they have to be individuated. It is not unlikely that this process is less efficient for the haptic modality. Besides being fast the subitizing mechanism is also characterised by accuracy. Regardless of the exact subitizing slope value, our data show that enumeration of small numerosities is much faster and more accurate than for larger numerosities. Furthermore, we have shown that this enumeration process is not the same as estimation. Therefore, we conclude that for small numbers of items (<4), regardless of relative differences between presented numerosities, an efficient mechanism analogous to the visual subitizing process is used. For larger numbers of items, either counting or estimation is used, depending on the size of the relative differences between the numbers of items.

An approach to the understanding of visual subitizing that is not based on discriminability, involves FINSTs (Fingers of Instantiation). This explanation is based on the Visual Indexing theory, which is the idea that humans have a way to refer to a certain item without having to link it to specific features of the item such as position (see Pylyshyn (2001) for a review). From multiple object tracking experiments it is known that subjects can track up to five items simultaneously (Pylyshyn & Storm, 1988). This leads to the hypothesis that there are five FINSTs that allow for up to five items to be tracked in parallel. It has been proposed that this also explains why small numerosities (<5) are enumerated more efficiently than larger numerosities (Trick & Pylyshyn, 1993, 1994). Visual experiments involving moving items have shown that also in that case subitizing can occur, even with the addition of distractor items (Alston & Humphreys, 2004). Note that in our haptic experiment the item positions were also not fixed. However, in this case the items could be physically manipulated by the subject and moved as the result of the subject's own action. This is of course never the case in visual experiments and physical item manipulation is specific for the haptic modality. It is possible that a haptic version of visual indexing exists or that visual indexing can also be used to process information that is extracted haptically.

Riggs et al. (2006) have reported evidence for the existence of an accurate and fast regime in haptic numerosity judgement for small numerosities (<4) when separate fingers were stimulated with pins. This is in agreement with the present study where we also find that there is evidence for two regimes in numerosity judgement. Gallace et al. (2006) have shown that there is no evidence for the existence of two regimes when vibrators distributed over the body surface had to be enumerated. In another study they have shown that also when vibrators were placed on the separate fingers results were quite similar to when they were distributed over the body surface and again they did not find an indication for two regimes in numerosity judgement (Gallace et al., 2008). However, for both modes of presentation (to the fingers and distributed over the body surface) error rates were very high (up to 90%). Already for a numerosity as small as two the error rate was 40% when the vibrators were presented to the fingers and from 4 items or larger the error rates are at chance level. This indicates that subjects could not assess the presented numerosities accurately in any part of the numerosity range. This suggests that perhaps vibrators are not suitable stimuli for this type of task. It is also possible that the presentation time (100 ms) was too short. Limiting presentation time is generally not desirable in numerosity judgement studies, since it may actually influence the enumeration

process (e.g. Jensen, Reese, & Reese, 1950; Trick, 2008). Short presentation time may force subjects, for instance, to use a faster but less accurate estimation process. Whatever the reason, the large error rates in the Gallace et al. study show that subjects could not enumerate the items correctly. Therefore, analysis of the accompanying response times does not provide insight into the enumeration process and it is not clear what should be concluded from this study.

The existence of two separate mechanisms (subitizing and counting) in visual numerosity judgement has been disputed. Balakrishnan and Ashby (1991, 1992) compared different models and showed that they could not find evidence for the existence of a discontinuity in numerosity judgement of small and large numbers of items. They concluded that enumeration was a continuous process and the upward bend is caused by an increasing cognitive load. One problem with this study is that they limited presentation time, which may influence the enumeration process as pointed out earlier. A more important point is that, for instance, an exponential function may approximate the shape of numerosity judgement data quite well. It is however not likely that response times will continue to increase exponentially with the number of items. Such a function can thus only be a good fit in the regime where it approaches a function with two linear parts. Therefore, finding a continuous function that fits the data does not rule out that there are actually two distinct underlying enumeration processes just as well as finding a good fit for a bilinear function does not necessarily mean that there are two distinct processes. Therefore, in Experiment 1 we compared performance of a bilinear function and a linear function to determine whether there was evidence for the existence of two regimes in the data. However, from these data alone it is not possible to conclude that there are two distinct enumeration processes. For drawing conclusions about whether separate enumeration mechanisms underlie performance in these regimes, one could manipulate the numerosity range and look at how performance changes. In vision this has, for instance, been done by Revkin et al. (2008). Therefore, we performed Experiments 2 and 3, which showed that when numerosity information is present, an enumeration process is used for small numerosities which is more efficient and accurate than estimation. This suggests that there exists an enumeration mechanism that is different from both counting and estimation in haptic numerosity judgement, similar to subitizing in vision.

Summarising, our results show that in haptic numerosity judgement there is evidence for the existence of two regimes in terms of response times and error rates. For small numerosities (<4) enumeration is more efficient and accurate than for larger numbers of items. Furthermore, we have shown that this efficient and accurate performance only occurs when numerosity information is present and does not depend on the relative differences between the numerosities in the range. It is unclear whether the haptic enumeration process is the same as the process underlying visual enumeration. Nonetheless, the pattern in the data is similar to that found in visual numerosity judgement studies. We therefore propose that the efficient enumeration of small numerosities in haptic numerosity judgement can be labelled haptic subitizing in analogy to the visual effect.

Acknowledgement

This research was supported by a grant from the Netherlands Organisation for Scientific Research.

References

- Akaike, H. (1974). New look at the statistical model identification. *IEEE Transactions on Automatic Control*, AC-19, 716–723.

- Akin, O., & Chase, W. (1978). Quantification of three-dimensional structures. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 397–410.
- Alston, L., & Humphreys, G. W. (2004). Subitization and attentional engagement by transient stimuli. *Spatial Vision*, 17, 17–50.
- Andres, M., Seron, X., & Olivier, E. (2007). Contribution of hand motor circuits to counting. *Journal of Cognitive Neuroscience*, 19, 563–576.
- Atkinson, J., Campbell, F. W., & Francis, M. R. (1976). The magic number 4 ± 0 : A new look at visual numerosity judgments. *Perception*, 5, 327–334.
- Balakrishnan, J. D., & Ashby, F. G. (1991). Is subitizing a unique numerical ability? *Perception and Psychophysics*, 50, 555–564.
- Balakrishnan, J. D., & Ashby, F. G. (1992). Subitizing: Magical numbers or mere superstition? *Psychological Research*, 54, 80–90.
- Camos, V., & Tillmann, B. (2008). Discontinuity in the enumeration of sequentially presented auditory and visual stimuli. *Cognition*, 107, 1135–1143.
- Carlson, R. A., Avraamides, M. N., Cary, M., & Strasberg, S. (2007). What do the hands externalize in simple arithmetic? *Journal of Experimental Psychology: Learning Memory and Cognition*, 33, 747–756.
- Cave, K. R., & Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology*, 22, 225–271.
- Coren, S. (1993). *The left-hander syndrome: The causes and consequences of left-handedness*. New York: Vintage Books.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433–458.
- Gallace, A., Tan, H. Z., & Spence, C. (2006). Numerosity judgments for tactile stimuli distributed over the body surface. *Perception*, 35, 247–266.
- Gallace, A., Tan, H. Z., & Spence, C. (2008). Can tactile stimuli be subitized? An unresolved controversy within the literature on numerosity judgments. *Perception*, 37, 782–800.
- Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience*, 6, 435–448.
- Jensen, E. M., Reese, E. P., & Reese, T. W. (1950). The subitizing and counting of visually presented fields of dots. *Journal of Psychology*, 30, 363–392.
- Kaufman, E., Lord, M., Reese, T., & Volkman, J. (1949). The discrimination of visual number. *American Journal of Psychology*, 62, 498–525.
- Lederman, S. J., & Klatzky, R. L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology*, 19, 342–368.
- Lederman, S. J., & Klatzky, R. L. (2004). Haptic identification of common objects: Effects of constraining the manual exploration process. *Perception and Psychophysics*, 66, 618–628.
- Logan, G. D., & Zbrodoff, N. J. (2003). Subitizing and similarity: Toward a pattern-matching theory of enumeration. *Psychonomic Bulletin and Review*, 10, 676–682.
- Luwel, K., Beem, A. L., Onghena, P., & Verschaffel, L. (2001). Using segmented linear regression models with unknown change points to analyze strategy shifts in cognitive tasks. *Behavior Research Methods, Instruments, and Computers*, 33, 470–478.
- Mandler, G., & Shebo, B. J. (1982). Subitizing: an analysis of its component processes. *Journal of Experimental Psychology: General*, 111, 1–22.
- Oyama, T., Kikuchi, T., & Ichihara, S. (1981). Span of attention, backward masking, and reaction time. *Perception and Psychophysics*, 29, 106–112.
- Pasini, M., & Tessari, A. (2001). Hemispheric specialization in quantification processes. *Psychological Research*, 65, 57–63.
- Pesenti, M., Thioux, M., Seron, X., & De Volder, A. (2000). Neuroanatomical substrates of arabic number processing, numerical comparison, and simple addition: A pet study. *Journal of Cognitive Neuroscience*, 12, 461–479.
- Piazza, M., Mechelli, A., Price, C. J., & Butterworth, B. (2006). Exact and approximate judgements of visual and auditory numerosity: An fMRI study. *Brain Research*, 1106, 177–188.
- Plaisier, M. A., Bergmann Tiest, W. M., & Kappers, A. M. L. (2008). Haptic search for spheres and cubes. In *Haptics: Perception, devices and scenarios*. In M. Ferre (Ed.). *Lecture notes on computer science* (Vol. 5024, pp. 275–282). Berlin/Heidelberg: Springer.
- Plaisier, M. A., Bergmann Tiest, W. M., & Kappers, A. M. L. (2009). Salient features in three-dimensional haptic shape perception. *Perception and Psychophysics*, 71(2), 421–430.
- Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80, 127–158.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3, 179–197.
- Reykin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, 19, 607–614.
- Riggs, K. J., Ferrand, L., Lancelin, D., Fryziel, L., Dumur, G., & Simpson, A. (2006). Subitizing in tactile perception. *Psychological Science*, 17, 271–272.
- Ross, J. (2003). Visual discrimination of number without counting. *Perception*, 32, 867–870.
- Ten Hoopen, G., & Vos, J. (1979). Effect of numerosity judgement of grouping of tones by auditory channels. *Perception and Psychophysics*, 26, 374–380.
- Theeuwes, J. (1993). Visual selective attention: A theoretical analysis. *Acta Psychologica*, 83, 93–154.
- Trick, L. M. (2008). More than superstition: Differential effects of featural heterogeneity and change on subitizing and counting. *Perception and Psychophysics*, 70, 743–760.
- Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 331–351.
- 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.
- Van Oeffelen, M. P., & Vos, P. G. (1982). A probabilistic model for the discrimination of visual number. *Perception and Psychophysics*, 32, 163–170.
- Watson, D. G., Maylor, E. A., & Bruce, L. A. M. (2007). The role of eye movements in subitizing and counting. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 1389–1399.
- Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, 9, 33–39.