Perception, 1976, volume 5, pages 327–334

The magic number 4 ± 0: A new look at visual numerosity judgements

Janette Atkinson, Fergus W Campbell, Marcus R Francis
Physiological Laboratory, University of Cambridge, Cambridge CB2 3EG, England
Received 6 April 1976

Abstract. Visual numerosity judgements were made for tachistoscopically presented linear arrays of dots or lines. The interelement interval (which could be specified in spatial frequency terms) was constant for a given array but varied across conditions. A clear limit in the accuracy of numerosity judgements was found to be set at 4 for regularly spaced elements with spatial frequencies below approximately 10 cycles/deg (element and interelement interval of 0·05 deg). This limit in terms of accuracy is accompanied by a fast and almost constant response time for arrays of 4 or less, compared to response times for arrays of more than 4 elements. The limit in accuracy falls to 2 elements rather than 4 for spacing narrower than 0·05 deg although with such spacing the elements are still easily resolved. The limit of 4 is found if the stimulus is a bright afterimage, lasting for approximately 60 s. This result suggests that the limit is independent of the time allowed for a single fixation and is a perceptual limit rather than a limit in some memory buffer. 'Numerosity' units are proposed to account for the results.

1 Introduction

In 1871 Jevons gave one of the earliest and clearest descriptions of an experiment on visual numerosity judgements. He says, "It is well known that the mind is unable through the eye to estimate any large number of objects without counting them successively. A small number, for instance 3 or 4, it can certainly comprehend and count by an instantaneous and apparently single act of mental attention". Jevons pointed out that the individual differences in instantaneous numerosity judgements may be due to perceptual grouping of the objects; obviously if objects are grouped into twos or threes you can comprehend as many groups as you can units, because, as he says, "the mind considers these groups only as units". When Jevons actually conducted the experiment on himself he avoided obvious perceptual groups and found that in "a single glance" the maximum number of items he could accurately estimate was 4.

Since 1871 an extensive body of psychological literature on the subject of numerosity judgements has developed (Glanville and Dallenbach 1929; Taves 1941; Saltzman and Garner 1948; Kaufman et al 1949; Jensen et al 1950; Klahr 1974). Nearly all of these studies have used either so-called 'random' arrays of dots, or dots arranged in a circle. In either case it seems that the arrays can be regarded as two-dimensional, and in the case of random arrays it is impossible to eliminate any possibilities concerning subjective perceptual groups. In the studies of Saltzman and Garner, Kaufman et al, Jensen et al, and Klahr tachistoscopic exposures varying from $\frac{1}{5}$ to $\frac{1}{2}$ s were used. Saltzman and Garner, Jensen et al, and Klahr also allowed unlimited time for estimation in some experimental conditions, thus allowing eye movements made during the exposure to either aid or hinder judgement of number.

Nearly all of the above studies have led to the conclusion that at least two processes are involved: one which holds for low numbers often called 'subitising' and an estimation process which operates for larger arrays.

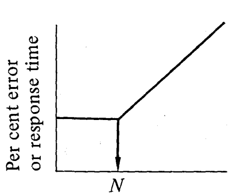
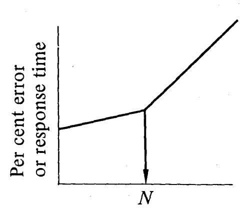
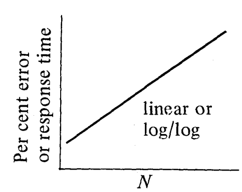
In fact three possible models have been suggested to account for the data which are shown in table 1, together with the results of previous experiments categorised under the various models.

The main problem in drawing any clear conclusions from these data is that there is much variety in the stimulus material used, the instructions and knowledge of results given to the subjects, and the time allowed for the task. For this reason we decided to use a simple spatial stimulus rather than a random or circular array, namely a line of short bars or dots. In all the experiments the elements were of high contrast (approximately 60%), i.e. well above contrast threshold. For such displays the spacing of the array can be clearly specified in terms of visual angle or spatial frequency of the fundamental, and the stimulus is spatially one-dimensional. For most of the experiments we restricted the exposure duration to 150 ms. Our reasoning for this was that most eye movements (except tremor) would not be possible in the exposure duration, so that only the information contained in a single glance, or single fixation, would be considered.

The second reason for choosing such a simple stimulus was that as much is known from neurophysiological data concerning spatial frequency and orientation channels, it might be possible to consider psychophysical results on numerosity judgement in the light of recent neurophysiological knowledge.

Recently Glezer et al (1973) have found units in the cat visual cortex which seem to operate similarly to a piecewise Fourier analysis, preferring as optimal stimuli black and white bars of a particular spatial frequency and particular number of cycles. It seems possible that these units may be related to similarly arranged units in human visual system, which could be optimally tuned to a particular number of cycles.

Table 1.

Model 1	Model 2	Model 3
		
Jevons (1871)—% error Saltzman and Garner (1948) —% error Kaufman et al (1949)—% error	Jensen et al (1950)—median response time Kaufman et al (1949)—median response time Klahr (1974)—response time	Kaufman et al (1949)—log/log median response time Taves (1941)—% error Saltzman and Garner (1948)—median response time

Note: Kaufman et al and Jensen et al both decided that the best straight-line fit to their data was the relationship between presented number and the reciprocal of the median response time minus a constant. Two different straight-line functions were fitted, a steeply negative one for $N = 1$ to 6 and a shallower slope for $N > 6$.

2 Experiment 1: the limit of accurate numerosity judgements

In the first experiment five observers were asked to report the number of dots in an oblique straight-line array of equispaced black dots on a white card (inset in figure 1); each black dot subtended 0·25 deg. The array was presented tachistoscopically with an exposure duration of 150 ms, which prevented the possibility of more than one fixation during the exposure. The centre of the array was centred on the fovea.

The dots were arranged in an oblique line which ranged randomly in angle from 30° to 45° either to left or right of vertical. The observer made a verbal response as soon as possible, and released a manual response key simultaneously with the verbal response. At least one practice trial was given on each of the cards used.

Figure 1 shows the mean response time of ten trials with each number of dots, N, for arrays containing 1 to 14 dots. The histogram shows the average percentage error on particular arrays. There is a small increase in response time for the arrays containing 1 to 4 dots. Two observers were shown the same stimuli, as in the previous experiment, but were now asked to respond only with "1", "2", "3", "4" or "more than 4". As can be seen from the results, also shown in figure 1, response times are very similar for all arrays except that containing 5 dots, where there is an additional delay of approximately 150 ms. The observer does not appear to know whether his response to 5 dots should be "4" or "more than 4". Errors in this task are only made for the array containing 5 dots (both observers made 20% errors on $N = 5$); for arrays of more than 5 dots the response "more than 4" is always correctly given. It seems from the results of these two experiments that no errors in numerosity judgements are made on arrays containing up to 4 dots, but that an array of 5 dots can sometimes be mistaken for 4 or sometimes 6 dots.

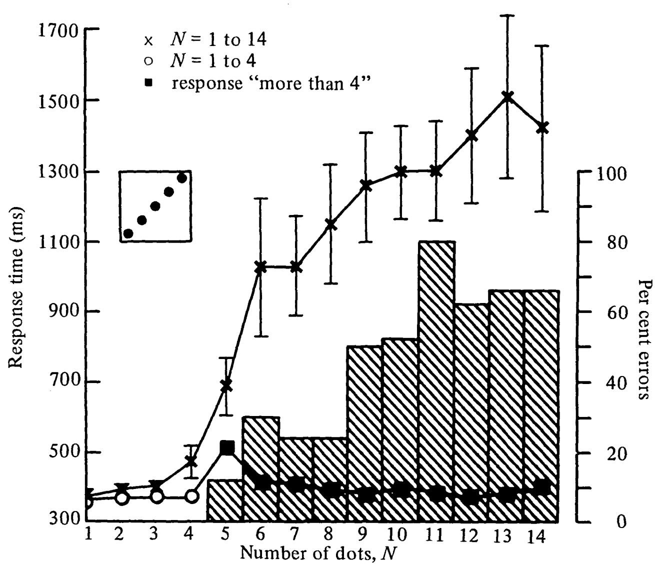


Figure 1. Mean response times and per cent errors as a function of the number of dots, N, for three different responses.

3 Experiment 2: the dependence of the limit in accurate numerosity judgements on spatial frequency

To see if the limit in accurate numerosity judgements is 4 when the spatial frequency is higher than 2 cycles/deg (i.e. the dots are closer together) two subjects were run on five different spatial frequencies. The fundamental spatial frequencies used were 2, 5, 8, 11, and 22 cycles/deg. The conditions were identical in other respects to experiment 1 except that the maximum number of dots was 10 rather than 14. Subject OJB was run on all the frequencies in a single session, the order of frequencies being randomised. Subject MF was run on 2 and 5 cycles/deg in one session, 8 cycles/deg in a different session, and 11 and 22 cycles/deg in a third separate session.

The results for the two subjects are shown in figures 2 and 3, and are quite similar for both subjects. In general up to 4 dots no errors are made at frequencies up to 8 cycles/deg. At 11 and 22 cycles/deg errors are made when N is 3 or more. The response times are relatively flat for $N = 1$ to 4 at lower frequencies, as was the case in experiment 1. At the higher frequencies there is an increase in response time for N greater than 2. The results suggest that the limit on accurate numerosity judgement is different for spatial frequencies below approximately 10 cycles/deg and those above this frequency. In the former case the limit is for $N = 4$, whereas in the latter case the limit is for $N = 2$.

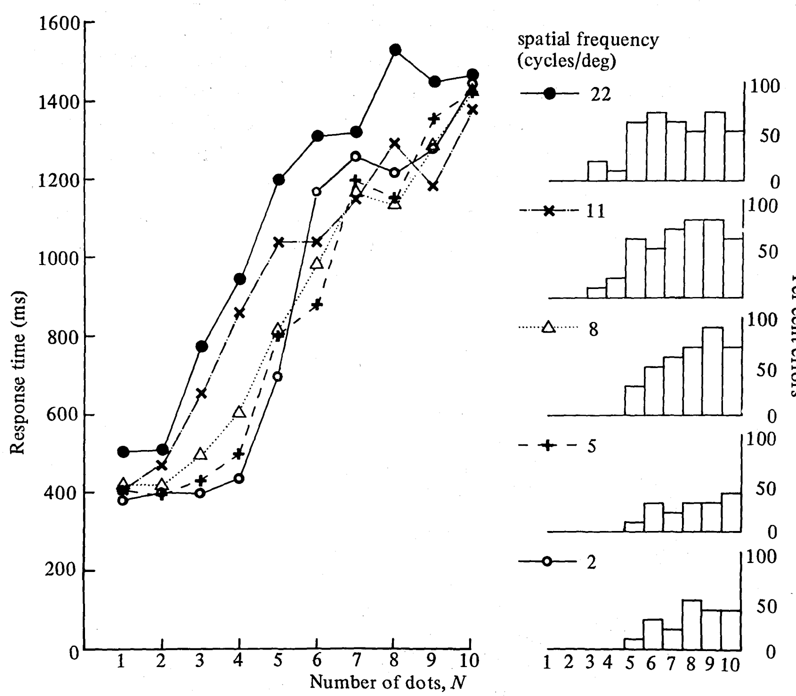


Figure 2. Mean response times and per cent errors for subject OJB as a function of fundamental spatial frequency of an array of dots.

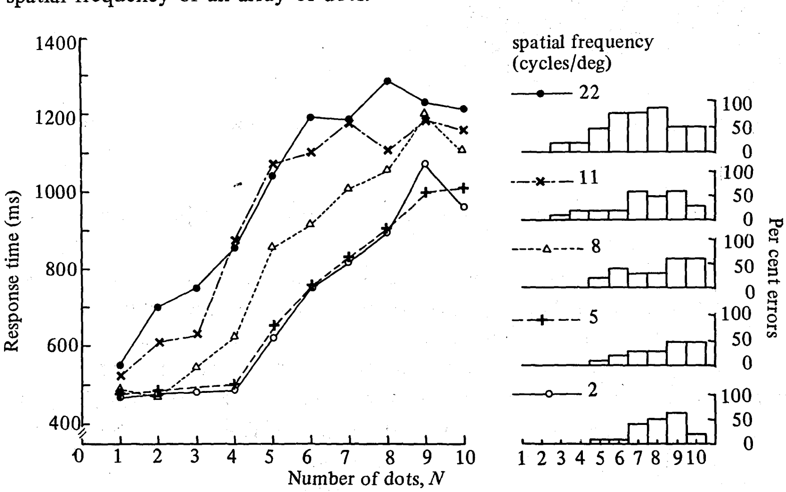


Figure 3. Mean response times and per cent errors for subject MRF as a function of fundamental spatial frequency of an array of dots.

How does this latter limit of 2 relate to the limit of resolving the space between dots at high spatial frequencies? In the 150 ms exposure the absolute threshold for seeing a space between two of the dots was approximately 25 cycles/deg, so that in the condition where the dots were at 22 cycles/deg the spacing was fairly close to the acuity limit. However, this is not true for the 11 cycles/deg condition in which counting was just as poor as the 22 cycles/deg condition. It seems that there is a range of spatial frequencies (11 to 25 cycles/deg) where the spaces between dots can be clearly resolved, but accurate counting above 2 is not possible.

To check whether this discrepancy between the limits of numerosity judgement and resolution was a function of the particular dot patterns used the same two subjects were tested at high frequencies for estimating the number of black lines exposed rather than black dots. For subject OJB this was in a separate session, whereas subject MRF was tested on the high spatial frequency dots and lines in the same session. The obliquely oriented lines, with square-wave profiles, were spaced at

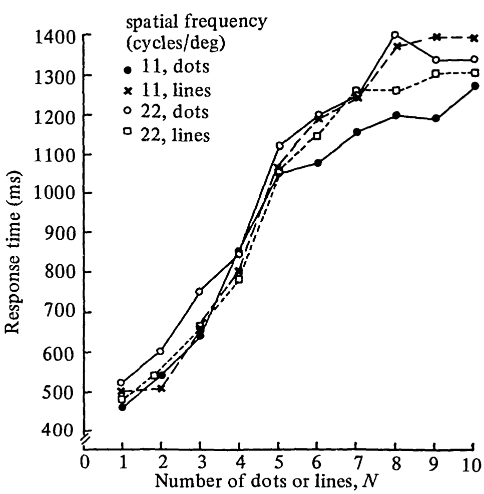


Figure 4. Mean response times for two subjects as a function of the number of dots or lines in an array, fundamental spatial frequency of the arrays being varied.

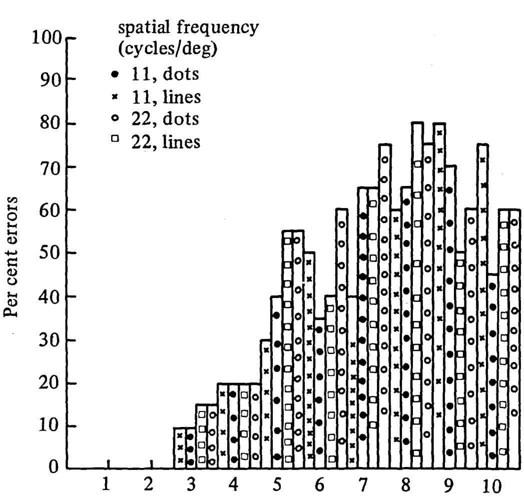


Figure 5. Mean per cent errors for two subjects as a function of the number of dots or lines in an array, fundamental spatial frequency of the arrays being varied.

either 11 or 22 cycles/deg. The resolution threshold for seeing a gap between two lines was much higher than that for two dots, being approximately 40 cycles/deg or 0·0125 deg resolution.

The mean response times for the four conditions are shown in figure 4 and the mean error histogram is shown in figure 5. Although response times are slightly longer for the 22 cycles/deg dot condition compared to the other three conditions, this difference failed to reach significance on analysis of variance. No systematic differences on the error data were found. Once again the accurate counting limit of $N = 2$ was found. The result suggests that the limit does not relate to the resolution values, but is related to the spacing or spatial frequencies of arrays. For spatial frequencies above approximately 10 cycles/deg, the number in arrays of more than 2 units cannot be accurately estimated, although the elements in the array are readily resolved as are the spaces between elements.

4 Experiment 3: the relationship between the numerosity limit and fixation time

The following experiment showed that even if the stimulus can be observed for 10 s, there is still a limit of 4 for arrays of low spatial frequency. Six observers looked at a bright afterimage of an array of dots (2·5 cycles/deg, obliquely oriented, see inset figure 6). The afterimage, produced by exposing the array in front of an electronic flash, was clearly seen for at least the first 10 s after the flash. Most observers still had a partial afterimage at the end of 60 s. Observers were asked to give a first estimate of the number of dots at the end of 10 s and a second estimate at 60 s after exposure. Mean accuracy on ten trials for each array is shown in figure 6. Again errors were made on arrays of 5 or more dots, the second estimates being slightly more accurate than the first. As the afterimage tends to partially fade and reappear during the course of 60 s after exposure, sometimes certain dots disappear while others remain visible. The observer used these extra cues from disappearances to improve his second estimate compared to his first estimate. However, errors on arrays of 5 dots are still sometimes made at the end of 60 s. The result suggests that it is not a short-term memory buffer that sets the numerosity limit.

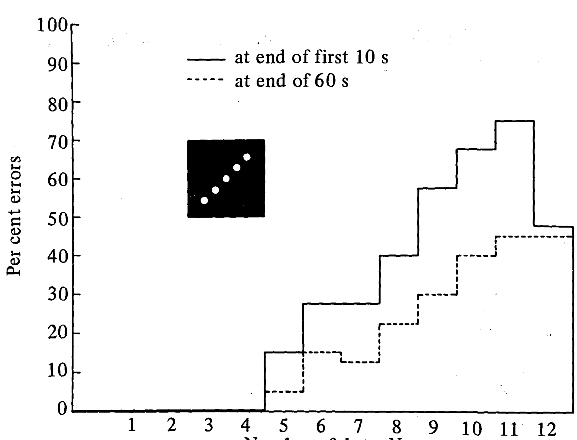


Figure 6. Mean per cent errors for six subjects as a function of the number of dots in an afterimage.

5 Discussion

The results of the previous experiments have shown that a clear limit in the accuracy of numerosity judgements is set at 4 for regularly spaced elements, with a fast response time accompanying this lower limit. This limit is dependent on the spacing of the elements in the array and holds for spatial frequencies below 10 cycles/deg (0·05 deg). The limit in accurate numerosity judgements drops to 2 elements for arrays where the elements are closer than 0·05 deg, although this is still well above resolution threshold. The limit would seem to be independent of the amount of time allowed for a single fixation of the array, although it is possible that the limit would be altered if eye movements were allowed, thus giving a number of fixations of the same array. The results suggest that some initial perceptual segregation process exists whereby the array is separated into visual chunks of a certain size.

Several hypothetical links can be postulated between these results and other psychophysical and neurophysiological results. One possibility is that the receptive fields of channels selectively sensitive to low spatial frequencies are in some way differently arranged compared to those selectively sensitive to high spatial frequencies. The former channels may have a number of excitatory and inhibitory side flanks which can readily respond to a number of different peaks and troughs of contrast in the stimulus, if appropriately spaced. Units of this type have been found by Glezer et al in cat cortex (1973). Two different sets of channels have been postulated from other neurophysiological (Enroth-Cugell and Robson 1966) and psychophysical (Kulikowski and Tolhurst 1973) results for high and low spatial frequencies; the present results on number estimation suggest that counting units are specific to the units tuned to low spatial frequencies. Yet another possibility has been put forward suggesting a link between the size of columns of neurones operating for different parts of the visual field and the ability to estimate the number of spaced objects (Barlow 1975). It might be that the channels that are used for numerosity are the same as those used to detect phase relationships; for both processes information as to the exact spatial location of two objects is needed for accurate perception. Such channels have been postulated previously by various investigators (Thomas 1970; Shapley and Tolhurst 1973; Kulikowski and King-Smith 1973; Nachmias and Weber 1975).

One possibility would seem to be that no matter wherever these units for counting are located in the visual system, they are the same units as those operating in tasks of perceptual segregation (Beck 1967; Olson and Attneave 1970) where an object must be located in visual space and segregated from objects and its background.

Acknowledgements. This research was supported by a Medical Research Council Grant (G 971/766/B). The authors would like to thank Professor Horace Barlow, Professor David Green and Dr Oliver Braddick for helpful discussions of the experiments reported.

References

Barlow H B, 1975 "Visual experience and cortical development" *Nature* **258** 199–204

Beck J, 1967 "Perceptual grouping produced by line figures" *Perception and Psychophysics* **2** 491–495

Enroth-Cugell C, Robson J G, 1966 "The contrast sensitivity of retinal ganglion cells in the cat" *Journal of Physiology* **187** 417–552

Glanville A D, Dallenbach K M, 1929 "The range of attention" *American Journal of Psychology* **41** 207–236

Glezer V D, Ivanoff V A, Tscherbach T A, 1973 "Investigation of complex and hypercomplex receptive fields of visual cortex as spatial frequency filters" *Vision Research* **13** 1875–1904

Jensen E M, Reese E P, Reese T W, 1950 "The subitizing and counting of visually presented fields of dots" *Journal of Psychology* **30** 363–372

Jevons W S, 1871 "The power of numerical discrimination" *Nature* **3** 281–282

Kaufman E L, Lord M W, Reese T W, Volkmann J, 1949 "The discrimination of visual number" *American Journal of Psychology* **62** 498–525

Klahr D, 1974 "A production system for counting, subitizing and adding" in *Visual Information Processing* Ed. W G Chase (New York: Academic Press) p 555

Kulikowski J J, King-Smith P E, 1973 "Spatial arrangement of line, edge and grating detectors revealed by subthreshold summation" *Vision Research* **13** 1455-1478
Kulikowski J J, Tolhurst D J, 1973 "Psychophysical evidence for sustained and transient detectors in human vision" *Journal of Physiology* **232** 149-162
Nachmias J, Weber A, 1975 "Discrimination of simple and complex gratings" *Vision Research* **15** 217-223
Olson R K, Attneave F, 1970 "What variables produce similarity grouping?" *American Journal of Psychology* **83** 1-21
Saltzman I J, Garner W R, 1948 "Reaction time as a measure of span of attention" *Journal of Psychology* **25** 227-241
Shapley R M, Tolhurst D J, 1973 "Edge detectors in human vision" *Journal of Physiology* **229** 165-183
Taves E H, 1941 "Two mechanisms for the perception of visual numerousness" *Archives of Psychology* **37** 1-47
Thomas J P, 1970 "Model of the function of receptive fields in human vision" *Psychological Review* **77** 121-134