Psychol Res (1982) 44:283–294

Psychological Research

What is Magic About the Magical Number Four?

Dietrich Simons and Dietrich Langheinrich

Psychologisches Institut der Universität Münster, Abt. Allgemeine Psychologie, Schlaunstr. 2, D–4400 Münster, Federal Republic of Germany

Summary. Previous studies have shown that the apprehension of number can be represented by three models according to the experimental procedure and the data analysis. The present experiment was designed to test the effect of figural characteristics of pattern on the response time. The subjects were asked to perform a same/different judgment, i.e., they were requested to decide whether a dot pattern, shown on a monitor, equalled a previously defined target number (n = 2–6) or not. Different 'types' of pattern were used and learning effects were studied. As was expected, the slopes for random and linear patterns were not so steep when the target number was low. With patterns in the dice mode, however, the slope was zero. Repeated presentations led to a slight reduction in slope for random and linear patterns only. In the case of the patterns in the dice mode, the repeated presentations caused only a change in the absolute reaction times (RTs) but had no effect on the slope. When the target numbers were larger (n = 5–6), the repeated presentations led to remarkable reductions in slope for random and linear patterns. The slope discontinuity at n = 4 occurred with all 'types' of pattern but it became less pronounced in the course of training at least in the case of random and linear patterns. This result is explained by clustering effects, use of figural cues, and a more efficient scanning process.

Introduction

The question how many different objects can be apprehended immediately dates back to Aristotle. Since then the question of the 'range of attention' has been discussed in various ways (Glanville and Dallenbach 1929). In the 19th century several researchers looked for a solution to the 'range of attention' problem by means of (quasi-) experiments (Hamilton 1859; Jevons 1871; Warren 1897). Since then the experimental procedure has hardly changed although it has become less and less clear whether the experiments really have anything to do with the original question or not.

Offprint requests to: Dr. Dietrich Simons, Parkstraße 5, D–5603 Wülfrath, Federal Republic of Germany

00340–0727/82/0044/0283/$ 02.40

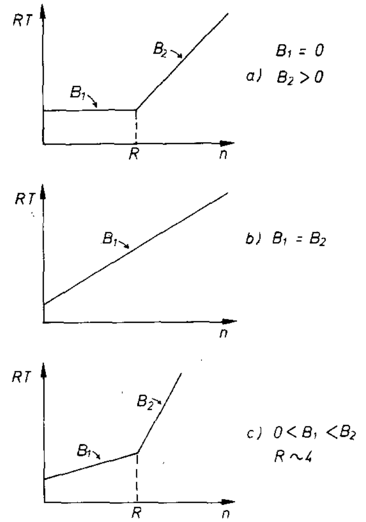

Fig. 1. Predicted reaction time as a function of *n*. R signifies a slope discontinuity at $n = 4$. After Klahr and Wallace (1976, p. 32)

We can distinguish between two different procedures (cf. Klahr and Wallace 1976). In both of them dot patterns are used as stimuli and the subject's task is to say how many dots he can see. In the threshold procedure (e.g., Averbach 1963) dot patterns are exposed for a brief period of time. 'The usual result of interest is the functional relation between exposure duration and the error rate for different values of n' (Klahr 1973, p. 7).

Model (a) in Fig. 1 represents the results commonly obtained by this procedure. In the latency procedure the dot pattern remains visible to the subject until he responds to the stimulus (e.g., Klahr 1973; Chi and Klahr 1975; Svenson and Sjöberg 1978). Experiments using the latter procedure have provided us among others with a new word called 'subitizing' (Kaufman et al. 1949). By this neologism the 'apprehension' of several dots 'in one single glance' is to be understood. 'Subitizing' occurs only with small numbers. Model (c) shows the basic differences between 'subitizing' and counting. Inspite of the fact that the mean RT increases with the number of dots presented there is a slope discontinuity at $n = 4$. Furthermore the slope for small numbers ($n = 1$–4) is different to that for larger numbers ($n = 5$–7). The range of 'subitizing' is determined by trend analysis, followed by linear regressions for both ranges (cf. Chi and Klahr 1975; Svenson and Sjöberg 1978). Using this kind of data analysis Chi and Klahr observed the following results: for $n = 1$–3: intercept = 495 ms, slope = 46 ms; for $n = 4$–7: intercept = - 442 ms, slope = 307 ms.

A synthesis of both procedures was suggested by Atkinson et al. (1976). They presented dots 'in an oblique line' with an exposure time of 150 ms in order to avoid eye-movements. The observer made a verbal response as soon as he could, and released a manual response key simultaneously with the verbal response (Atkinson et al. 1976, p. 329). Atkinson et al. observed very nearly identical RTs (about 400 ms) for the range of numbers $n = 1$–3. There was a definite change in slope between $n = 4$ (about 460 ms) and $n = 5$ (about 700 ms). The only errors occurred when the pattern was made up of five or more dots. The same was true when a number of white dots on a black background were presented with the aid of an electronic flash in such as way that an after-image lasting at least 10 s was produced. Atkinson et al. believe the numerosity limit at $n = 4$ to be a 'perceptual limit rather than a limit in some memory buffer,' based on 'numerosity limits' (p. 327).

Model (b) is less important. Results which corresponded with this model were found, for example, by Saltzman and Garner (1948), who compared the median response time to the number of dots presented.

The reaction time is not determined only by the number of dots. In addition to this it may be affected by figural characteristics of the pattern (Beckwith and Restle 1966; Schaeffer et al. 1974), by eye-movements and the visual angle (Klahr 1973). The inter-individual differences (cf. Svenson and Sjöberg 1978) lead us to suggest, however, that the subjects use quite different strategies when performing their task. The analysis of Svenson's and Sjöberg's data (for the numbers $n = 1-3$) reveals a rather high correlation of $r = 0.70$ between intercept and slope (N = 16). The results may also depend on the 'type' of pattern used in the experiment: Klahr (1973) preferred random patterns, Svenson and Sjöberg presented dots in a horizontal row, and Atkinson et al. presented the dots in an oblique line, but for technical reasons only. The experiment reported here is based primarily upon Klahr's approach, but we use a different method:

Procedure

Instead of asking the subjects how many dots they can see ('identification'), a same/different judgment is required of them, i.e., whenever a previously defined number is visible on the monitor the subject has to answer 'Yes,' otherwise 'No.' Whereas in the traditional approach the subject is shown a differing number of dots in random order, the method used here enables the subject to 'concentrate' on a particular number of dots. It is supposed that the subject forms an ever-decreasing number of pattern 'types' (clustering) in order to make it easier for himself to make a match between the given number of dots and those presented on the visual display. This should lead to a decrease in the slope per dot particularly when the number of dots is higher than $n = 4$.

Stimulus material

In addition to random and linear patterns which have been used as stimulus material in most research into 'apprehension of number,' this study introduces further pattern 'types' in order to ascertain the effect of pattern familiarity on the RT for each successive number of dots. It is expected that the use of familiar patterns such as those found on dice will lead to a reduction in slope because the subjects will be able to draw upon figural cues.

Training

Saltzman and Garner (1948) reported that repeated presentation can lead to a reduction in RT. A comparison of the trials No 1, No 2, and No 10 of altogether 10 trials shows a marked reduction in RT for numbers consisting of more than four dots. The improvements occurred primarily during the first five trials. It was supposed that the method used here of presenting particular number of dots 'en bloc' would favor learning processes as, for instance, the subjects would be paying attention to figural cues. This supposition was to be tested in an analysis of learning.

Subjects, Materials, and Apparatus

Subjects. Six subjects from the psychology department participated as volunteers.

Materials. The materials consisted of patterns of one to eight small rectangles (7 x 6 mm) which were called 'dots'. The dots were presented on a video monitor (controlled by a TRS 80) within a visible frame of 9.5 cm^2. The dot patterns did not exceed a projection area of 8 cm^2. Each number of dots (apart from 1) was made up of different patterns drawn up as follows. Both figurative and mixed patterns were shown to the subjects. Figurative patterns could be found in the form of a 'closed' pattern (e.g., a pattern in the dice mode) as well as in an 'open' form (e.g., linear patterns). One pattern was chosen at random. In addition to this, the size and position of some of the figurative patterns was varied. Figure 2 illustrates the pattern when the target number is $n = 5$. The number of different pattern variations for each value of n are as follows: $n = 1$ (1), $n = 2$ (6), $n = 3$ (10), $n = 4$ (18), $n = 5$ (19), $n = 6$ (24), $n = 7$ (29), $n = 8$ (40).

The target number was $n = 2$–6. The target numbers were combined in two series (1: n=2,3,5; 2: n=3,4,6) for technical reasons. The presentation of three dots as target number in both series made possible an analysis of the effect of the numbers that come before and after the target number on the RT. The order in which the target numbers were to be presented (in the series) was determined before the start of each series, e.g., 4-3-6. Then all patterns containing four dots (=18) were put in random order and presented in a ratio of 1 (target) : 2 (non-target) patterns. The distractors were chosen from

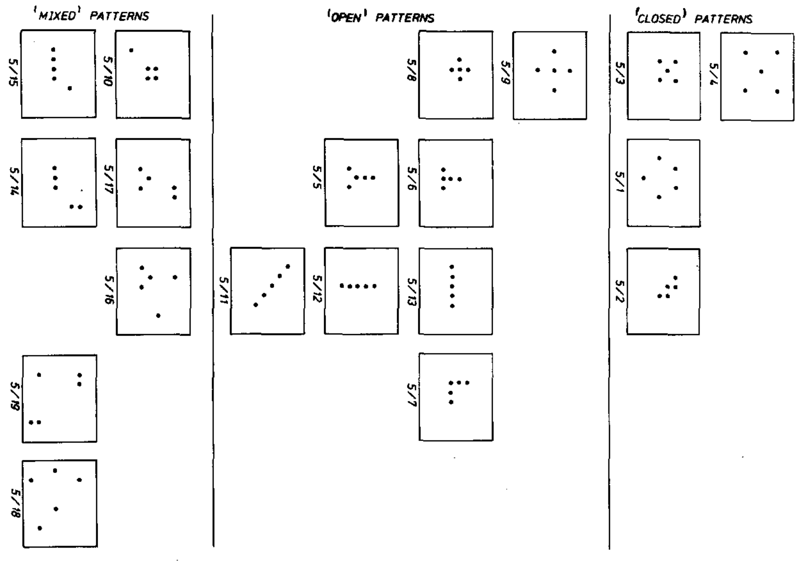

Fig. 2. Five-dot pattern, subdivided into three 'types' of pattern. 'Closed' patterns change their 'Gestalt' when further dots are added. The opposite is true for 'open' patterns. 'Mixed' patterns consist of common 'open' or 'closed' patterns plus additional dots or pairs of dots

and equally distributed among the remaining numbers (n=1–3, 5–8). As each pattern of a target number was only shown once, the number of trials per target depends on how many patterns there are for each value of n. In the case of n = 4, there are 18 different possible patterns plus 2 x 18 distractors = 54 trials. Series 1 contained 105 trials, series 2, however, 156 trials. Series 1 took 8 min to complete and series 2 11 min. There was a break of about 10 min between each series.

Procedure and Design

The subjects were seated 130 cm[2)] in front of the video monitor, placed at eye level. They were requested to keep their hands lying on the keys during one series of trials. The following procedure is shown in Fig. 3. At the beginning of each series the subjects were told which target number was going to be tested. Before changing the target number the subjects were requested to concentrate on the next number. By pressing both keys simultaneously one series of trials was started. First the frame was generated, then two very small rectangles, starting from the vertical line of the frame moved to the center of the projection area and disappeared. This procedure was aimed at helping the subjects to pay attention at the right moment and to watch the center of the projection area. The instant the two rectangles met each other, the screen was blacked out and the pattern generated. Following this the pattern was shown to the subject. By pressing one of the two keys the subjects had to decide whether the number of dots visible on the monitor corresponded to the target number previously determined ('Yes') or not ('No'). When, for example, the subject was requested to watch the fours (= target number) he had to press the 'Yes' key whenever he saw four-dot patterns.

The whole experiment consisted of six sessions on six consecutive days. Both series (n=2,3,5; n=3,4,6) were shown alternately during a session. The order of presentation

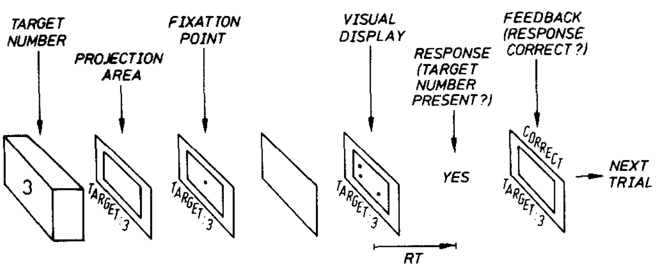

Fig. 3. The experimental procedure: 1) the target number is shown; 2) the projection area is terminated; 3) the fixation dot comes on; 4) the dot pattern is generated (the monitor turns black for a short moment); 5) the dot pattern is presented; 6) by presenting one of the two keys the subject decides whether the dot pattern matches the target number; 7) feedback is given

---

2) A pre-test, using the same methods described here, showed that the visual angle had no systematic influence on the average RT. As a result of Klahr's (1973) data, which shows that the spread of the pattern can be important, we came to the conclusion that the viewing distance is also important. The distance of 130 cm was subjectively judged to be particularly pleasant

of the target numbers within a series was balanced out over all six sessions. In order to balance out serial effects during a session, the experiment was planned for six subjects. Although the experiment began with eight subjects, four subjects were found to have incomplete data so that the quantitative evaluation is based on the other four.

Results

The error rate was 5.3% for target numbers and 3.7% for non-target numbers. The error rate for the target numbers $n = 2$–4 was 3.4%, the rate for $n = 5$–6 was, however, 7.2% ($P < 0.05$). The error rate for small numbers is different to the rate reported by Chi and Klahr: $n = 1$–3: 0.87% ($n = 4$–9: 8.9%). The error rates were the same for all six sessions. Of the non-target errors 85% were committed when the non-target number did not exceed the target number by more than one dot (Table 1). When $n = 4$ is the target number, the likelihood of a false alarm is far greater when the non-target number is $n = 3$ than when the non-target number is $n = 5$, i.e., the 'quantitative distance' between '3' and '4' is apparently smaller than that between '4' and '5.' The number of errors rises quite clearly when $n = 5$ or $n = 6$ are the target numbers, especially when the non-target numbers are one more than the target numbers.

Table 1. Mean reaction times and error rates (with percentage put in parentheses) for every number (in italics) or non-target number and the number of trials per number

Number of dots								Number of trials per number	
1	2	3	4	5	6	7	8	Target	Non-target
400 (0)	*354 (7)*	343 (0)	322 (0)	301 (0)	282 (2)	287 (0)	282 (0)	144	41
357 (4)	446 (2)	*372 (2)*	400 (0)	345 (0)	332 (0)	301 (0)	296 (0)	240	69
330 (0)	376 (0)	451 (10)	*415 (3)*	472 (1)	379 (1)	370 (0)	336 (0)	432	123
345 (0)	347 (0)	373 (2)	522 (8)	*518 (6)*	600 (18)	483 (2)	413 (0)	456	130
345 (0)	349 (0)	368 (0)	441 (0)	602 (15)	*600 (8)*	728 (30)	576 (8)	576	165

Table 1 also reveals an interesting distribution of the RTs for non-target numbers: the mean RT decreases in relation to the quantitative difference between target and non-target numbers. In this case it is easier to decide on: 'considerably more than the target number' than 'considerably less. . .' It is remarkable that for target numbers $n = 2$–4 more time is needed for non-target numbers that are one less than the actual target number than for the target number itself.

The analysis of learning effects reveals significant decreases in RT during the first two sessions.[3] A Friedman test was performed, the results for all six sessions are as

[3] This result was already apparent in the pre-test

follows: Chi_r^2 (10.24, $P = 0.05$) : $n = 2$: 2.64; $n = 3$: 12.46; $n = 4$: 8.68; $n = 5$: 10.29; $n = 6$: 11.85. The analysis of the last four sessions did not reveal a significant decrease in RT for any number. A rather peculiar result in the relatively small decrease in RT for number $n = 4$ compared to $n = 3$.

As has already been mentioned, the target number $n = 3$ was presented as target in both series, in other words, twice as often as the other target numbers. This procedure made research possible into three further questions, related to learning analysis. The results are as follows: 1) It is of no importance to the average RT whether the target number $n = 3$ is present in the series $n = 2, 3, 5$ or $n = 3, 4, 6$. 2) Repeating the target number $n = 3$ during a session leads to the same result; and 3) although the target number $n = 3$ was target number 12 times instead of the usual 6 times there was no improvement in the average RT after the sixth run through. Consequently performances are not affected by previous or subsequent target numbers or by short-term repetitions.

Table 2 shows the results of linear regression for some experiments with unlimited exposure time. The calculation of the RT slopes in this experiment are based on the means per target number (for all patterns, sessions, and subjects). For n=2–5 both a linear and quadratic trend were significant (F(4.75) = 51.8 linear, 6.27 quadratic). Therefore a linear regression was performed for $n = 2$–4 and $n = 4$–6. The results are: $n = 2$–4: intercept = 290 ms, slope = 31 ms; $n = 4$–6: intercept = 32 ms, slope = 97 ms. Although Chi and Klahr (1975) discovered, so to speak, two slope discontinuities, namely from $n = 3$ to $n = 4$ and from $n = 4$ to $n = 5$, in this experiment there is only one slope discontinuity at exactly $n = 4$. The predicted values for the target number $n = 4$ are very close to the empirical value: 415 ms vs. 414 ms ($n = 2$–4) or 420 ms ($n = 4$–6). It is remarkable that the slope for $n = 2$–4, but especially for $n = 4$–6, is considerably lower than those in the reported studies in which random patterns were presented to the subjects.

Table 3 illustrates the average RT for random and linear patterns and those in the dice mode.[4] First let us consider the general means (sessions 1–6). Linear patterns are

Table 2. Results of linear regression in experiments with unlimited exposure time

Author	Stimulus range	Analysis range	Intercept (ms)	Slope (ms/dot)	Pattern
Klahr (1973)	1– 5	1– 4	421	66	random
	1–10	1– 4	451	72	random
		6–10	-114	268	random
	1– 5	1– 4	324	25	r. (no eye movem.)
		1– 4	308	60	r. (more than one fixation)
Chi and Klahr (1975)	1–10	1– 3	459	46	random
		4– 7	-442	307	random
Mandler and Shebo (1982)	1–20	4–20	-348	382	random
S. and L.	2– 6	2– 4	290	31	different 'types'
		4– 6	32	97	of pattern

[4]Because of the number of subjects it was decided to dispense with the analysis of variance

Table 3. Mean reaction time (ms) as a function of n for different 'types' of pattern. The mean for the first two sessions (1–2) is compared with the mean for the last four sessions (3–6). The mean for all six sessions (1–6) is also represented

	Sessions	Number of dots 2	3	4	5	6
Random	1–2	387	456	499	672	841
patterns	3–6	363	427	451	527	580
	1–6	371	437	467	575	667
Dice mode	1–2	400	382	396	451	469
patterns	3–6	339	334	341	371	424
	1–6	359	350	359	400	439
Linear	1–2	302	381	437	662	955
patterns	3–6	330	363	421	504	538
	1–6	321	369	426	557	677

easier than random patterns when n = 2–4, but approach random patterns when n = 5–6. At n = 2–4 both linear and random patterns demonstrate the expected slope of roughly 50 ms; the slope at n = 4–6 with 100–130 ms lies clearly below the values that Klahr reported. 2) Patterns in the dice mode are, relatively speaking, the easiest. When n = 2–4 they show a slope of zero, at n = 4–6, however, the slope is roughly 40 ms.

If one takes the analysis of learning into consideration and compares the RTs of the learning phase (sessions 1–2) with those of the last four sessions, then one begins to see a slightly different picture. Random and linear patterns show a considerable decrease in RTs for larger target numbers (n = 5–6). Random patterns show a decrease of about 30 ms when the target numbers are smaller (n = 2–4); linear patterns show no continuous change. In the last four sessions both the random and the linear patterns show a more linear development for all target numbers (n = 2–6), i.e., the slope discontinuity at n = 4 is declining.

Patterns in the dice mode show a mean decrease in RT of about 60 ms for all target numbers. So during the learning phase the RT improvement is highest for those patterns that are most familiar to the subject. Despite the remarkable improvements, the slopes for the target numbers n = 2–4 as well as those for the target number n = 4–6 remain constant.

As was mentioned earlier (see Table 2) 'mixed' patterns were also shown to the subjects. The general question was, how the RTs for 'mixed' patterns would compare with those for figurative patterns. Within the framework of this experiment the question can be defined more clearly. If one adds a single dot or pair of dots to a basic figurative pattern (examples: Table 2, pattern 5.10, 5.17), how do the different RTs for the basic pattern and the one that has been increased by one or two dots relate to one another? This analysis was carried out with seven basic patterns and is illustrated in Fig. 4. Each value for x is based on 24 single values. It is notable that if a five-dot pattern is increased by one dot, then the difference in RT between the basic pattern and the increased pattern is about the same as when one increases a three- or four-dot pattern by a pair of dots. The three- or four-dot patterns have a comparative time-difference

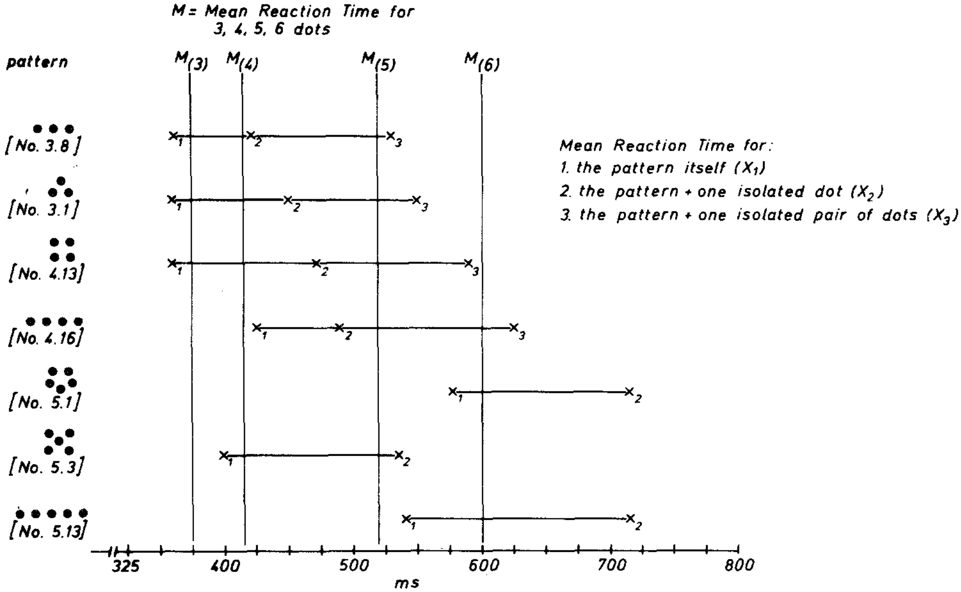

Fig. 4. Mean reaction times for some 'closed' and 'open' patterns and 'mixed' patterns composed of them. The overall mean reaction times for the numbers $n = 3$–6 are represented by vertical lines

when they are increased by a pair of dots. Despite some results to the contrary one gains the impression that the differences in RT between two patterns containing the same number of dots remain the same when more dots are added. The results suggest at least that the groups of basic patterns plus isolated dots have been taken over by the subjects, i.e., the subjects do not perceive such patterns as one single 'Gestalt' (see Simons and Langheinrich, Ref. Note 1).

Discussion

This experiment has made three things clear: 1) The same/different procedure leads to slopes that are less steep, especially when the target number is more than four dots even when the pattern is random. 2) Familiar patterns (those in the dice mode for example) have a slope of zero for target numbers $n = 2$–4. 3) The absolute RT for all 'types' of pattern decreases with training. In the case of linear and random patterns, the slope decreases to such an extent by larger numbers ($n = 5$–6), that the slope discontinuity at $n = 4$ is not very distinct. In the case of the patterns in the dice mode the slope discontinuity at $n = 4$ conflicts with Mandler and Shebo's (1982)[5] results. Mandler and Shebo also showed their subjects patterns in the dice mode (canonical patterns) at a fixed exposure time (200 ms) and let them give their response by way

[5]Mandler and Shebo's study was not published until after this experiment had been completed.

of a voice key. The RTs for the numbers $n = 1-5$ were all identical (600 ms), the RT for $n = 6$ was about 900 ms. What do these results mean for the understanding of the slope discontinuity at $n = 4$ or even $n = 5$?

Patterns of three, or even of four, dots do not look 'structureless' even though they are called 'random' patterns. What we see are pairs, lines, triangles, and rectangles (cf. Mandler and Shebo 1982). The number $n = 3$ can be represented by four distinct patterns, the number $n = 4$ by about eight (except 'mixed' patterns). On the basis of a simple match in STM one might therefore expect equal RTs for the numbers $n = 1-4$, but that is only true for patterns in the dice mode, i.e., those patterns with which we are very familiar. This is also Mandler and Shebo's argument. They maintain that canonical patterns are learned during the course of one's life: they become equivalent to 'chunks.' There is no automatic apperception of numerosity, rather an inference that a triangle is made up of three dots. In conformity with our results this can only be said of patterns in the dice mode. It is still unclear why there is a slope discontinuity at $n = 4$ (in our experiment) or $n = 5$ (Mandler and Shebo) even for patterns in the dice mode. Equal RTs for the target numbers $n = 2-4$ are obviously the exception to the rule.

The basis for the apprehension of number is a scanning process (cf. Klahr 1973) which takes about 40 ms per item. Patterns of not more than four dots are seldom 'structureless' on the one hand and they are not very complex or complicated on the other. The scanning process of larger numbers becomes more difficult insofar as it needs more control to avoid scanning one dot twice or to avoid missing one dot. Larger numbers can be represented by a variety of patterns some of which are hardly organized at all ('random') and some of which are very 'geometrical' ('pentagon', 'hexagon', etc.). In both cases it may be a problem where to start and where to stop the scanning process: it is a problem of "unpacking". In accordance with Beckwith and Restle (1966) and Schaeffer et al. (1974), Klahr and Wallace argue that

> stimuli are represented as a cluster, or group, of elements that is first detected at a global level and then 'unpacked'. 'Unpacking' can be viewed as a detailed attending act in which the gross features of the visual field are examined more precisely (1976, p. 44).

As already mentioned, some 'geometrical' ('pregnant') patterns offered 'resistance' to the 'unpacking'. One of the most difficult six-dot patterns showed a dot in the middle of a regular 'pentagon'. Only two out of 24 patterns yielded longer RTs. Subjects found this pattern very difficult in the beginning but less so after training. The decrease in RT is surely due to an improved 'unpacking' as well as to improved figural cues. Especially when seeing a pattern for the first time, the subjects reported that they had to make every effort to concentrate on the dots and not on the figure itself (e.g., 'cross'). Proceeding from the assumption that patterns in the dice mode are so common that there is nothing more to learn about them, the mean decrease in RT (about 60 ms) during training may be due to a faster response selection. (By subtracting 60 ms from the total decrease in RT one might get the time per pattern which represents the different processing times before and after training).

The remarkable decrease in RT during training occurred especially with the numbers $n = 5-6$. The slopes for patterns in the dice mode (about 40 ms from the very beginning) and for linear and random patterns (about 60 ms) might suggest that the variety of patterns ($n = 5: 19; n = 6: 24$) was reduced by clustering and that the scanning process was improved. Whether these results are limited to the same/different judgment is open to discussion. The slope discontinuity at $n = 4$, i.e., the magic number four plus or minus zero (Atkinson et al. 1976), seems to represent a 'law', provided that the subjects are not familiar with the patterns. Model (c) seems to be the rule, but by presentation of rather common, geometrical patterns or by training, the slopes, predicted by the models (a) or (b), may be confirmed.

Finally it should be asked whether the subjects used a more general strategy when performing the task. In 'detection theory' terms one could argue that the speed of 'Yes' answers depends solely on how familiar the target pattern is to the subject. The speed of 'No' answers depends not only on how familiar the target pattern is but also on the numerical difference between target and non-target. In the case of a numerical difference of more than three dots the subjects need only to decide between 'much more than the target number' or 'much less.' The RTs for those non-target numbers are consequently very fast. Not only the RT but also the amount of errors made increases when the non-target numbers are close to the target numbers, and that in turn depends on the number of dots. The fact that considerably fewer mistakes are committed when the non-target number is $n = 5$ in comparison with the non-target number $n = 3$ when $n = 4$ is the target number could be an argument in favor of the qualitative difference between $n = 4$ and $n = 5$. The data for $n = 5$–6 support the supposition that the subjects followed the rule of thumb 'in dubio pro target' when the non-target number contained one more dot than the target number. On the other hand, in the case of the non-target patterns that contained one dot less, the subjects were more likely to check the pattern again to make sure that they had not missed a dot. One could say that the subjects tried to avoid missing the target number. They obviously preferred a false alarm to a miss. Moreover it took less time to see that a dot pattern contained one more dot than the target number than one dot less.

Reference Notes

1) Simons D, Langheinrich D: Die Bedeutung figuraler Strukturen für die Mengenerfassung. Gestalt Theory, accepted for publication.

References

Atkinson J, Campbell FW, Francis MR (1976) The magic number 4 ± 0: A new look at visual numerosity judgements. Perception 5:327–334

Averbach E (1963) The span of apprehension as a function of exposure duration. J Verb Learn Behav 2:60–64

Beckwith M, Restle F (1966) Process of enumeration. Psychol Rev 73:437–444

Chi MTC, Klahr D (1975) Span and rate of apprehension in children and adults. J Child Psychol 19:434–439

Glanville D, Dallenbach KM (1929) The range of attention. Am J Psychol 41:207–236

Hamilton W (1859) Lectures on Metaphysics and Logic. Mausel HL, Veitch J (eds) Lecture 253. Blackwood, Edinburgh

Jevons WS (1871) The power of numerical discrimination. Nature 3:281–282
Kaufman EL, Lord MW, Reese TW, Volkman J (1949) The discrimination of visual number. Am J Psychol 62:498–525
Klahr D (1973) Quantification processes. In: Chase WG (ed) Visual Information Processing. Academic Press, New York
Klahr D, Wallace JG (1976) Cognitive Development. Lawrence Erlbaum Associates, Hillsdale, NJ
Mandler G, Shebo BJ (1982) Subitizing: An analysis of its component processes. J Exp Psych Gen 111:1–21
Saltzman IJ, Garner WR (1948) Reaction time as a measure of span of attention. J Psychol 25:227–241
Schaeffer B, Eggleston VH, Scott JL (1974) Number development in young children. Cognit Psychol 6:357–379
Svenson O, Sjöberg K (1978) Subitizing and counting processes in young children. Scand J Psychol 19:247–250
Warren HS (1897) The reaction time of counting. Psychol Rev 4:569–591

Received February 24, 1982/June 30, 1982