CHAPTER 2 Meaningfulness and frequency

In this chapter all numbers 1 through 100 are put to a test of continued association[1]. In such a setting, people generate as many associations as they can think of for each stimulus, within a limited amount of time. The assumption is that in performing this task, subjects set up an abstract representation of the stimulus. This representation links the stimulus word (or number) to other words (or numbers), which are subsequently produced as responses. Representations containing more conceptual links will allow people to produce more associations. As an implicit knowledge test it can be used to identify different levels of ability.

The continued association test serves several purposes here. The first is to determine m-scores for each of the individual numbers in our study. The m-variable (Noble, 1963) is a measure of a number's productivity as a stimulus, which is taken to be an indication of the richness of its representation. In word studies, m has been found to distinguish between concrete and abstract words, the former obtaining higher m-scores (De Groot, 1989). In the case of numbers the abstract-concrete distinction is not quite appropriate. However, other objective (that is, non-mental) distinctions may serve a similar purpose. One of these is size. In general, large numbers are known less well that small numbers (cf. Paulos, 1989, for a complaint about this mental limitation). Moreover, strong size effects are invariably found in studies of mental arithmetic, suggesting that, on the average, larger numbers are more difficult to process. Another distinctive property is membership of the multiplication tables. Of the stimulus numbers, 53 are members of one or more of the multiplication tables (the numbers 1-12 and their products) and 47 are not. The large category of tabled numbers may be further divided into subcategories, for instance separating single-digit numbers and tens from the others. The first purpose of the experiment is, in sum, to determine if and how m distinguishes between numbers.

The second purpose is to further investigate and classify associatively produced relationships between numbers. In this respect the experiment bears some resemblance to that performed by Costermans (1990b), who asked his subjects to produce three associations to each number from 1 to 20. Costermans subsequently identified four main classes of relationships, being neighbourhood (N+1), symmetry relative to 10 (e.g., N-10 or 10 - N), simple ratio's (N/2) and common membership of some typical number families, such as primes.

The present experiment is different in two respects: it studies a larger sample of numbers, and sets no limit (other than time) on the number of associations to be produced.
In Chapter 1 it was mentioned that children use different criteria for conceptual grouping than adults. For instance, young children will judge the numbers 7 and 8 to be more alike, because they are neighbours. Older children and adults, in contrast, prefer to group numbers according to attributes such as evenness, or membership of the same multiplication table. Thus, they will judge 4 and 8 to be more like rather than 7 and 8 (Miller & Gelman, 1983; Shepard et al., 1975; see also Costermans, 1990a). Such attributes, however, may not be associatively available for all numbers. In a previous study with 10 and 11 year old children, we found that multiplicative relationships can be more or less transparent (Milikowski, 1987). A ratio of 30 : 10, for instance, was better understood and more ably used than a ratio of 54 : 18. It is possible that such differences are also present in adults, and will be expressed in an association task.
The third purpose of this chapter is to obtain an approximation of relative frequencies for the set of numbers participating in this study. That frequency is an important dimension of knowledge is generally accepted, though its role in word association has been the subject of some debate. As mentioned in Chapter 1, reaction times in mental calculation are a function of problem-size: the larger the two operands of an addition or a multiplication problem are, the slower subjects come up with the answer. This is the case even within the set of numberfacts (addition and multiplication of single-digit numbers), where answers seem to be more or less automatically retrieved from memory (Ashcraft, 1982,1992; Ashcraft & Battaglia, 1978; Campbell & Graham, 1985; Parkman, 1972; Parkman & Groen, 1971). Several authors have recently proposed that such RT-differences could be an effect of frequency rather than size. This explanation is supported by the negative correlation between size and frequency that is seemingly universal (Dehaene & Mehler, 1992), and by the fact that textbooks used in primary education contain many more small-sized than large-sized problems (Hamann & Ashcraft, 1986; Ashcraft and Christy, 1991, cited in Ashcraft, 1992). In their article comparing number-word frequencies in several languages, Dehaene and Mehler do in fact state that the effects of problem size "might as well be attributed to word frequency" (Dehaene & Mehler, 1992, p. 22). A similar (though not identical) line of reasoning is offered by Ashcraft: "Just as frequently occurring words enjoy a privilege in lexical and semantic access, frequent problems enjoy an advantage in the memory system for arithmetic" (Ashcraft, 1992, p. 102).
However, the relationship between stimulus frequency and speed of responding is not as straightforward an issue as these authors suggest. Though it is true that frequency was once viewed as a reliable predictor of response times (frequent words being responded to more quickly), De Groot (1989) has demonstrated that this view may have been based on a confounding of frequency with concreteness (c.q. imageability). In her experiments, in which frequency and concreteness were manipulated independently, stimulus frequency had no effect whatsoever on response times, while concreteness had. A similar pattern was

found for the prediction of m: the critical variable was concreteness, while frequency as such had no significant effect.

If De Groot's findings apply to numbers as well as words, the explanation of the size effect as an effect of frequency would not hold. To investigate this point (which will be addressed in a further chapter) we need frequency scores for numbers. The problem is that general language frequencies are only available for a limited set of (frequently used) numbers, such as single-digit numbers, teens and tens. Of the hundred numbers to be studied here, 72 have unknown frequencies. This brings us back to the third objective of the present chapter. One of the results of an association experiment is a set of responses, in this case a set of response-numbers. These numbers can be ordered and scored for the frequency of their occurrence. The representiveness of the obtained distribution can subsequently be tested by comparing, for the 28 numbers with known frequencies, the scores obtained from counts of general language usage with the experimentally obtained scores. If these scores are sufficiently well correlated, we could use the associative frequency measure as an approximation of the relative frequencies produced in general language settings.

Of course, this treatment of frequency is not without problems. I will discuss these in more detail after the results have been presented. There is one possible objection, however, which I will briefly comment on here. In psychology, word frequency is traditionally treated as an input-variable, and not as a measure of output. It must be observed, however, that frequency is at least an output-measure also. Frequency-data are gathered by counting the words that are actually produced in books and newspapers. The measure in fact represents the grand sum of usage across contexts. The production of subjects in an association experiment is similarly decontextualized, be it by a different method. Thus viewed, it is not so strange to expect similarities between the frequency distributions within these sets.

Method

Stimulus categories

All numbers 1-100 were used as stimuli. For purposes of comparison, the set was divided into two main categories by separating tabled numbers (n = 53), from non-tabled numbers (n = 47). Some subcategories will also be compared. Tabled numbers can be subdivided into single-digit numbers (n = 9), tens (n = 10), and other tabled numbers (n = 34). The category of non-tabled numbers can be subdivided into primes, which have no divisors, and other non-tabled numbers, which have divisors. Table 2.1 gives the members of each category.

Table 2.1

Five categories of numbers and their exemplars.

Category	***Stimuli***
Tabled numbers (53)	

Single-digit numbers (9)	1, 2, 3, 4, 5, 6, 7, 8, 9
Tens (10)	10, 20, 30, 40, 50, 60, 70, 80, 90, 100
Other tabled numbers (34)	11, 12, 14, 15, 16, 18, 21, 22, 24, 25, 27, 28, 32, 33, 35, 36, 42, 44, 45, 48, 49, 54, 55, 56, 63, 64, 66, 72, 77, 81, 84, 88, 96, 99
Non-tabled numbers	
Primes (20)	13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97
Other non-tabled numbers (27)	26, 34, 38, 39, 46, 51, 52, 57, 58, 62, 65, 68, 69, 74, 75, 76, 78, 82, 85, 86, 87 91, 92, 93, 94, 95, 98

Subjects

Subjects in the experiment were 18 first year psychology students fulfilling a course requirement. Subjects were tested in a group. The duration of the experiment was about an hour.

Materials

The materials consisted of booklets with a separate page for each number. All booklets contained all numbers 1-100. The order in which the stimuli were presented was different for each subject. A stimulus number was printed at the top of each page. It was also printed at the beginning of each line, to discourage chaining responses. Such chains are produced when people associate to the previously given response, instead of to the stimulus itself.

Procedure

Testing took place in a classroom. After the subjects were seated, the experimenter read out the instruction, which was adopted from Noble (1963), with two modifications. Both stimuli and responses were to be numbers instead of words, and subjects would be given 30 instead of 60 seconds per stimulus for their responses. As in Noble's instruction, chaining responses were explicitly warned against, and subjects were advised to think back to the stimulus number each time before writing down a new response. After the instruction was read out, the booklets were distributed. Following a sign from the experimenter, the subjects started on their first number. Time was kept with a stop-watch. After fifty numbers a short break was taken.

Results

M-scores and stimulus categories

Examples of subjects' performance in this task are given in Table 2.4. M-scores were computed by counting the associations produced for each stimulus number (see Appendix I) and averaging these total scores over subjects. The mean m-score for all numbers is 8.06, which means that slightly more than eight associations were produced per number by the average subject. The twelve highest and the twelve lowest scoring numbers are presented in Table 2.2. The highest scoring numbers are 12 and 50, with an average of ten associations. The lowest scoring numbers are 67 and 76, with average m's between six and seven. A one way analysis of variance was performed, comparing the means of the two main stimulus categories, those of tabled and non-tabled numbers. The differences are statistically significant, $F(1, 98) = 30.85$, $p < .0001$.

Table 2.2

Highest and lowest scoring numbers on *m*, in order of rank (12 ranking highest and 67 lowest).

Top 12	***Bottom 12***
12	53
50	77
10	74
60	87
15	78
16	98
100	43
11	59
34	68
22	85
69	76
93	67

Mean category and subcategory values are given in Table 2.3. Tens obtained the highest m-scores (8.92), followed by the subcategories of single-digit numbers (8.41), and other tabled numbers (8.22). Prime numbers above twelve and other non-tabled numbers each had significantly lower m-scores (7.62, and 7.73 respectively). A post hoc comparison between the subcategories revealed that each tabled subcategory scored reliably better than each non-tabled subcategory ($p < .05$). The other differences between subcategories failed to reach significance.

Table 2.3

Mean *m*-scores of five categories of numbers. Standard deviations are given within brackets.

Category	***m-score***	***(st. dev.)***
Tabled numbers (53)	8.39	(.66)
Single-digit numbers (9)	8.41	(.32)
Tens (10)	8.92	(.70)
0ther tabled numbers (34)	8.22	(.64)

Non-tabled numbers (47)	7.67	(.60)
Primes (21)	7.62	(.56)
Other non-tabled numbers (26)	7.73	(.63)
Total (100)	8.06	(.72)

Size and *m*

Numerical size is negatively correlated with m (r = -.37). This suggests that the contrast between tabled an non-tabled numbers might be partly an effect of size. In a multiple regression analysis using size and tabledness as predictors, a multiple R of .53 was obtained. Both variables make a significant contribution to the prediction of m. The standard regression coefficient (beta weight) of tabledness is .41, p < .0001. The beta of size is –.22, p < .05. Thus, tabledness is the stronger predictor.

Table 2.4

Examples of two subjects' response patterns in continued association. The examples contain the complete set of associations produced by subjects F and L to the stimulus numbers 9, 12, 13, 50 and 67.

Stimulus	***Subject F***	***Subject L***
9	3, 2, 1, 18, 90, 99, 33, 6, 36, 54, 27, 81, 900	
12	1, 2, 21, 6, 3, 120, 24, 144, 36, 60, 72, 1200, 1212, 212	
13	1, 3, 10, 30, 31, 26, 169, 130, 310, 260	
50	5, 10, 500, 0, 25, 100, 150, 5000, 51	
67	6, 7, 134, 68, 60, 70, 9	[illegible], 113

Associative relationships between numbers

93, 60, 1, 2, 3, 14

4, 9, 83, 51

5, 500, 5000, 3, 9, 12

For all hundred stimulus numbers, a response hierarchy was determined. Leading responses (also called dominant, or primary, responses), being those named by the greatest proportion of subjects, were then scored for their relationship to the stimulus number. All leading responses are listed in Appendix 2. They fall into four categories. Mean frequencies of these categories are presented in Table 2.5. The table also gives the mean commonality of each category, indicating the proportion of subjects who produced a response which belongs that particular category. The largest category (42 cases) consists of divisor- and multiple-responses. Examples of divisor-responses are 6 : 3 and 14 : 7. Examples of multiple-responses are 10 : 100 and 45 : 90. A second category (18 cases) contains the digit-responses. Here, the subjects responded with one of the digits of a multi-digit numbers. Examples are 43 : 3 and 58 : 8. This category contains both the "base ten-responses" identified by Costermans (1990b), and a second group of responses in which the tens are treated as a single digit (e.g., 53 : 5). The third main category (with 27 cases), is named ambiguous. This category contains all responses that can be interpreted in more ways than one. Most responses in this category are both divisors and digits. Examples are 77 : 7, and 15 : 5.

Table 2.5

Main categories of leading (most frequent) responses, with their frequency of occurrence in column 2, and their mean commonality (proportion of subjects giving the particular response in column 4. For each response category, two examples are given in column 3..

Category	*N*	*Examples*	*Commonality*
Divisors (31) & multiples (11)	42	64: 8; 10:100	65%
Digits	18	43 : 3; 58 : 8	48%
Double meanings	27	77 : 7; 15 : 5	63%
Others	13	19:20; 67:13	52%
Total	100		58%

The fourth category, named others, contains such leading combinations as 67: 13 (the response is the sum of the stimulus number's digits), 19 : 20 and 59 : 60 (the response is the stimulus number's next neighbour) and 79 : 97 (the two digits have changed places). For one stimulus number (78) two responses occurred with equal frequency (8 and 87).

Each leading response has its commonality-score, representing the proportion of subjects who named that response. Since people could write down all associations which occurred to them during half a minute, commonality scores are relatively high. The stimulus-response pair 10 : 100 even obtained a commonality-score of hundred percent (see Appendix 2). This percentage means that each subject has named 100 as an association to 10.

Two observations seem to indicate that subjects tend to avoid digit- responses. The first is that most leading associations belong to other response categories. The second is that, if a digit-response is the leading one, it is backed by a relatively small proportion of subjects. An inspection of the leading associations given in Appendix 2 shows that digit-responses are exclusively associated with non-tabled numbers. Tabled numbers have no leading digit-relationships at all. The leading associations of these numbers are either divisors or multiples, or belong to the ambiguous category. That all leading digit-associations occur in the categories of non-tabled numbers may indicate that such a response is only given when no other options seem available.

Frequency distribution of responses

Of all 14503 responses, 11876 (82 %) are whole numbers between one and a hundred.

Response frequencies were determined for each of these numbers (see Appendix 1). Correlations were then computed between those response frequencies and four of the general language frequency measures used by Dehaene and Mehler (1992).[2] As mentioned in the introduction to this chapter, general language frequency counts are available for 28 numerals only, being the single-digit numerals 1-9, the teens (10-19), and the other multiples of ten, including 100. Several language counts do not even provide data points for each of these 28 numbers. We will limit our comparison to those that do (see Table 2.6). They are American English, Spanish, Dutch (numerical), and Dutch (all instances). The first three counts only contain cardinal and ordinal numerals, excluding, for example, the adverbial meaning of "second" in English, and meanings as in "she has no second". The second Dutch count (all instances) includes such forms of adverbial usage. The four measures are logarithmic transformations of the absolute scores. It should be noted that all four concern number words, not Arabic digits.

Table 2.6.

Correlations between size and several frequency variables for 28 numbers.

	1	2	3	4	5	6	7
1 Size	1						
2 Spanish (all)	-.24	1					
3 Dutch (all)	-.41	.91	1				
4 American (num.)	-.62	.84	.91	1			
5 Dutch (num.)	-.43	.85	.91	.93	1		
6 Associative frequency	-.57	.81	.82	.88	.89	1	
7 Log associative frequency	-.57	.78	.80	.86	.88	.98	1

Note. Of the correlated frequency counts, the first two (all) are based on all applications of a number-word, while in the other two (num.) the count was restricted to numerical meanings only.

Table 2.6 gives the correlation matrix for six variables: 1. Size. 2. Spanish (all instances) 3. Dutch (all instances) 4. American (numerical) 5. Dutch (numerical) and 6. Associative frequency, as determined in the present experiment.

Strong correlations were obtained between associative frequency as measured in the present experiment, and the four general language frequency counts (see Table 2.6). The correlations are .81 (Spanish, all instances); .82 (Dutch, all instances); .88 (American, numerical) and .89 (Dutch, numerical). Taking the logarithm of the associative frequencies does not improve these correlations (see Table 2.6, row 7). Therefore, the absolute frequencies will be used for further comparisons.

Table 2.7

Highest and lowest scoring numbers on associative frequency, in order of rank (number 2 ranking highes,t and number 79 lowest).

Top 12	*Bottom 12*
2	82
3	78
1	94
4	57
6	59
8	53
9	87
7	71
5	85
11	73
10	67
12	79

Table 2.7, which gives the numbers scoring highest and lowest on the frequency variable, may be compared with the selections based on m-scores (Table 2.2). There is more overlap on the negative poles (53, 78, 59, 85, and 67) than on the positive (12, 10, and 11).

Categories, *m* , and associative frequency

A one way analysis of variance was performed on the frequency scores of the categories and subcategories that were also compared for differences in m (see Table 2.3). The overall comparison between tabled and non-tabled numbers was significant, $F(1, 98) = 46.41$, $p < .0001$). The frequency differences between the subcategories was very large. Table 2.8 gives the mean frequencies, which can be compared with the mean m-scores given in Table 2.3.

Table 2.8.

Mean associative frequency-scores of two categories and five subcategories of numbers. Standard deviations are given within brackets.

Category	*Ass. frequency*	*(st. dev.)*
Tabled numbers (53)	10.00	(7.08)
Single-digit numbers (9)	22.96	(4.11)
Tens (10)	10.59	(2.86)

Other tabled numbers (34)	6.47	(3.63)
Non-tabled numbers (47)	2.78	(1.68)
Primes (20)	3.13	(2.33)
Other non-tabled numbers (27)	2.51	(0.92)
Total (100)	6.60	(6.38)

The subcategory of single-digit numbers has a mean frequency score of 22.96. Tens, with a mean score of 10.59, also score well above the grand mean of 6.60. Prime numbers from thirteen up and other non-tabled numbers have mean frequencies of 3.13 and 2.51 respectively. Except for the one between prime and other non-tabled numbers, all comparisons are statistically reliable ($p < .05$).

The difference between the categories are much larger for frequency-scores than for m-scores. This is well illustrated by Figure 2.1. M measures a number's associative productivity, whereas associative frequency measures the same number's chance of occurring as a response. These measures are clearly not symmetrical, as an inspection of Figure 2.1 will confirm. M-scores show a relatively even distribution between the (sub)categories, while the distribution of associative frequency scores is highly uneven. Single-digit numbers, and to a lesser degree tens, are used disproportionately often as an associative response, while prime and other non-tabled numbers are hardly ever named.

Maar waar is de tabel gebleven?

Figure 2.1.

M-scores and Associative frequency scores of five subcategories of numbers.

Size effects

Within the set of 28 numbers for which general language frequencies are available (see Table 2.6), the correlation between size and associative frequency was .57. For the complete set of hundred numbers, the correlation between size and associative frequency was .65. The correlation between m and associative frequency was .50. When the category variable (tabled versus non-tabled) and the variable size were entered in multiple regression, the multiple R with associative frequency was .85. Standard regression coefficients were -.62 for tabledness, and -.34 for size. Both are significant, $p < .0001$.

A comparison of numbers and words

We have seen that the distribution of associative-frequency scores is much more uneven than the distribution of m-scores. This is well illustrated by Figure 2.1. Numbers with high frequency scores (single-digit numbers and tens) are produced as responses to many other numbers. In contrast, numbers with low frequency scores (all non-tabled numbers) are only produced in specific contexts.

A similar asymmetry may govern the response distribution in word association. We have tested this notion by ordering the response words obtained by De Groot (1981) according to their frequency of occurrence. The response list given by De Groot contains all words named in an association experiment, together with their sources (stimulus words) and the frequencies of response-stimulus combinations. In De Groot's study, 100 subjects produced one single association to each of 460 stimulus words. Thus, 46000 responses were obtained. A count (performed by hand) learns that around 5070 different words are involved, which gives an average response frequency of nine per stimulus word. These figures cannot directly be compared with the frequency scores in the present experiment, in which the conditions were different. It is possible, however, to compare frequency distributions. Interestingly, the distribution of De Groot's response set is very similar to our own. It is also highly uneven. While many words occur only once, certain other words occur very often. Twelve words have scores higher than 250. These twelve words, which represent 0.2% of all named words (12 out of 5070), account for 9.5 % of all responses (4391 out of 46000). The list is interesting, because it seems to reflect some primary human concerns. The leading response words are, in order of (experimental) frequency :

1. Water (904); 2. Food (469); 3. Bird (401); 4. Tree (379); 5. Sea (337); 6. Animal (300); 7. Horse (291); 8. Pain (275); 9. Coat (263); 10. Shoe (259); 11. Cow (252); 12. Fire (251).

These twelve words stand for things of great importance in human life. Their high occurrence in the associative response set indicates that mental organisation is sensitive to such importance. Of course, the response list may be biased to a certain degree by the selection of stimuli. Nonetheless it offers an interesting illustration. The comparison seems to confirm that, for words and numbers alike, differences between concepts as measured by associative frequency of being used are much larger than those measured by m. More speculatively, one might observe that frequency, as a variable, seems to be related to ecological significance.

Discussion

As stated in the introduction, the experiment reported in this chapter had three objectives. The first was to collect m-scores for all numbers 1-100, and to compare the m-scores of the different (sub)categories of numbers. The second was to classify associative relationships between numbers. The third was to obtain an approximation of relative frequencies within the set of 100 numbers. In this section I will first sum up the results. Then I will offer some interpretations, discuss some problems and state the main conclusions.

All numbers 1-100 were introduced as stimuli in an experiment using continued association. Five categories were compared: 1. Single-digit numbers, 2. Tens, 3. Other tabled numbers, 4. Primes (from thirteen up), and 5. Other non-tabled numbers (see Table 2.1 for a complete classification). The three first-named subcategories were each found to have higher m-scores than the two last-named. This difference between the main groups of tabled versus non-tabled numbers is significant, though not very large. No significant differences were obtained between the tabled subcategories. Neither was a significant difference obtained between the subcategories of non-tabled numbers.

Most leading responses were found to belong to the following categories: 1. Divisors and multiples (42%), 2. Digit responses (18%), and 3. Ambiguous, the response being both a divisor and one of the digits of the stimulus number (27%). Pure digit-responses have lower commonality-scores than the other two classes.

associative frequency was found to correlate between .81 and .89 with general language frequencies for 28 (relatively frequent) numbers. A weaker correlation, of –.65, was obtained for associative frequency and objective size. Single-digit numbers and tens were produced as an association disproportionately often. A comparison with a response distribution obtained by Annette de Groot indicates that a similar asymmetry may govern the associative production of words. I will now discuss some implications of these outcomes.

In the first place, it is evident that non-mental attributes of numbers can be used to predict mental performance with a reasonable degree of precision. In general, subjects' performance was better for the tabled numbers than it was for the non-tabled numbers . In this respect all three measures were found to converge. The different categories of numbers have different response patterns. For some numbers, many types of responses seem to be available, which can be illustrated by the difference between 16 and 61. These two numbers are composed of the same digits. However, 16 has meaningful associations with 8 and 4, which in this experiment are the preferred responses. Divisors are not available for 61, which is a prime and has 6 (a digit-response) as its leading association. For some smaller prime numbers, such as for 13, the preferred response is a multiple. For 61, however, no multiples seem to be associatively available. Neither, for that matter, are divisors in the case of many divisible non-tabled numbers, such as 78. This shows that associative responses do not simply reflect "objective" numerical attributes. Such objective attributes are part of some number-concepts, but not of others. Both components (objective and subjective availability) influence the m-variable.

At the same time, it seems clear that m captures one aspect only of a concept's psychological meaning. This is most evident for single-digit numbers. Considered as stimuli, these numbers are not exceptionally meaningful, producing an average of 8.41 associations, as against 7.60 for the lowest (primes) and 8.92 for the highest scoring subcategory (tens). This moderate m-score of single-digit numbers stands in sharp contrast to the very high frequency of the category's occurrence as a response. This frequency is 22.96, while the lowest scoring numbers (other non-tabled) have an average frequency of 2.51 only. The comparison makes clear that numbers which do not generate many associations by themselves, may nonetheless be named as associations to other numbers very often. It implies that, psychologically viewed, relationships within the domain are not symmetrical. Larger numbers, it seems, are primarily understood by people in terms of smaller numbers. One probable cause is that these smaller numbers do in fact serve many functions. They can, for example, be a divisor of the stimulus number, one of its digit-components, or both.

A small number contributes to the psychological meaning of several lager numbers, while the reverse relation is often absent. Thus, a number such as 5 is generated as an association to many of its multiples, but few multiples are generated as an association to 5. One way to view this phenomenon is suggested by Barsalou (1989, 1990, 1993), who makes a difference between context-dependent and context-independent elements in the representation of concepts. Taking 5 as an for example, frequently occurring associations such as 10, 25, 50 and 6 might be considered to be part of its conceptual core (context-independent), while numbers to which 5 is only related as a response, such as 95 or 65, would be part of its context-specific fringe (context-dependent). However, this explanation is relevant to one aspect of the data only. It does not address the differences in frequency of context dependent usage. While single-digit numbers and tens are part of the conceptual core of many others, most numbers are denied such an important role.
Similar mental asymmetries have been studied by Rosch (1975), who discriminated between "reference point concepts" and other concepts within the same category. Rosch used these terms to describe a certain hierarchy in the relationship between concepts. In one experiment, Rosch' subjects were requested to fill in sentences such as "… is essentially (or roughly / basically / almost, etc.) an …". Reference point stimuli - which in the case of numbers were thought to be multiples of ten - were usually placed in the second slot. In a second experiment subjects placed a given stimulus in a spatial relationship to an other stimulus, whose position was fixed. It was found that the mental distance between the two types of concepts was asymmetrical. When a reference stimulus was fixed, the other stimuli were placed close to it. In the reverse condition, when a reference concept must be manipulated in relation to a non-reference concept, larger distances were produced. For instance, 79 would be placed closer to 80 than 80 to 79. Similar results were obtained for other categories, such as colours and lines. Focal colours and vertical or horizontal lines were found to function as reference point stimuli.

It is evident that psychological attributes of numbers can be investigated by means of word association techniques and word association variables. There are some reasons, however, why the present results cannot directly be compared with those obtained in a classical free word association experiment. In such experiments the stimulus words are taken from various categories, and any word may be given as a response. In our experiment both stimuli and responses belonged to the category of numbers. Moreover, the formal numbers system is neat and economical. It contains no redundancy in the form of synonyms, and its syntax is simple: Ten digits suffice to make all numbers. Each of these differences may, in its own way, have contributed to the constrained response set, which has a 82 percent overlap with the set of stimulus numbers. Such a degree of overlap will never be obtained in word association.

However, interesting similarities were also discovered. In the Results section, the frequency differences observed in the present experiment were compared with those obtained in a word association experiment. Response frequencies is both experiments are highly uneven, indicating a hierarchy in associative relationships Of particular interest is that the most frequent response words all referred to necessities of human life. The most frequent association was water, being followed by food. In our experiment, single-digit numbers were the most frequent responses. It seems reasonable to assume that these numbers have greater psychological importance than larger ones.

Main findings

¶ Different categories of numbers have different psychological properties. The psychological variables m and associative frequency have moderately strong correlations with objective numerical properties, such as size and membership of the multiplication tables. Associative frequency is strongly correlated with general language frequency.

¶ Associative frequency and m describe two different dimensions of a number's psychological meaning. M is a stimulus variable, while associative frequency is a response variable. These two variables, which are also correlated, have different distributions, indicating an asymmetry in the mental associations between numbers.

¶ Single-digit numbers have very high associative frequencies, suggesting a privileged mental status. Certain words seem to be similarly favoured. The most frequent responses in an experiment of De Groot were water and food, which stand for things of great importance.

1 Some of the results reported in Chapter 2 have also been published elsewhere. See Elshout and Milikowski, 1990; Milikowski and Elshout, 1994.

2 The data used for this analysis were kindly provided by Stanislas Dehaene.