CHAPTER 4 Liked and disliked numbers

Introduction

The subject of this chapter is people's private - or at least subjective - feelings about numbers.[1] There is obviously more to the meaning of numbers than either a formal description or a frequency count will reveal. Some people associate numbers with colours (Seron et al, 1992). Many people have favourites and lucky numbers. Some people also have special dislikes, which may or may not be superstitious. Hotels often skip 13 as a room number, and sometimes as a complete floor, because their clients may suffer from tristaidekafobia, which is Greek for the fear of thirteen (Wells, 1987). The encyclopaedia further informs us that eleven, "going out above ten", stands for madness, sin and intemperance, while three, seven and nine have magic qualities.

Some subjective associations may be accidental. A person who was born on the 5th of May seems predestined to like 5, especially in the Netherlands, which was liberated on that date in 1945. Such a connection does not depend on any intrinsic property of 5 except its smallness; any number up to 12 might come up as a happy date. The Armistice which ended World War I was declared on the eleventh of November, the eleventh month, but there is nothing in 11, beyond its regular occurrence as a day and as a month, to explain that particular connection. Because the knowledge of such dates is shared by many people it may, in certain contexts, come up as a part of a number's representation as produced by a large group. This cannot be the case with associations that are private only, such as birth dates and house numbers. Since counting starts at 1 and must necessarily end somewhere, a number's chances of occurring in any such context will be largely determined by its size (cf. Dehaene, 1992; Dehaene & Mehler, 1992).

Other subjective associations may, however, be more specifically related to individual properties of numbers. As Hans Ebersark once wrote in a lyrical description of the joys of mental calculation: "I wouldn't go as far as some of my fellow calculators and indiscriminately welcome all numbers with open arms: not the horny-handed, rough-and-tough bully 8 or the sinister 64 or the arrogant, smug, self-satisfied 36. But I do admit to a very personal affection for the ingenious, adventurous 26, the magic, versatile 7, the helpful 37, the fatherly, reliable (if somewhat stodgy) 76..." (Smith, 1983, p. Xlll). Such opinions, though

private and subjective, are nonetheless based on objective properties of the numbers concerned. Of course, this is an expert's opinion. Most people have never experienced 37 as particularly "helpful" (see the comments on that number printed in Table 4.1), because they neither share Eberhart's "consuming passion for numbers", nor his "ability to handle them, fondle them and bring them to heel" (Smith, 1983, p. XI). But most people have some expertise in handling numbers. Some properties of some numbers are well entrenched in common knowledge. Educated adults have spent several thousands of hours calculating, practising and developing their knowledge of numbers. This may result in certain feelings and opinions. To investigate whether such opinions are indeed held, and how they relate to numbers' individual and common attributes, was the purpose of the experiments to be reported in this chapter. The body of the chapter is made up by two experiments. In the first, all numbers 1-100 are rated on twelve scales of the Semantic Differential (Osgood, Suci & Tannenbaum, 1957). In the second, opinions about numbers are studied by means of a questionnaire asking for favourite numbers, hated numbers and examples of other subjective categories.

Prior to the two experiments we interviewed fifteen people about their preferences and dislikes concerning numbers. The results of this pilot study are rather interesting in themselves, and can be read as an introduction to the main experiments. Each interview had two parts. In the first, subjects were presented with some numbers and asked for their comments: what can you tell me about this number? Tell me anything that comes to your mind. An illustration of the comments collected on four different numbers is presented in Table 4.1.

Table 4.1

A selection of subjects' comments on four different numbers.

25	
Subject 1	It is divisible only by 5. So it is not really… agile
Subject 5	It makes me think of a quarter. 25 cents.
Subject 8	It is the square of 5. A beautiful number.
Subject 9	That is my age. A quarter of a century.
37	
Subject 1	That is a number you can't do anything with. It means nothing to me, 37. I have nothing to say about it.
Subject 5	That's the age of my boyfriend.
Subject 6	It means you are settled in life. At that age, you have got things worked out a bit.
Subject 8	It is prime. That's all.
Subject 9	It means nothing to me.
Subject 12	I can make nothing of it.

Subject 15 It does not appear in any of the multiplication tables. And that is the only math I know.

2

Subject 1 That is my favourite number. You can make a duck out of it and - well to me it is a very beautiful number. 1 is lonely and 2 is much nicer. It means company. And it is an even number, which is important to me, because I don't think odd numbers are very interesting. Even numbers are much more interesting. They are nicer to look at, and to write down.

Subject 5 A pair. Two people.

Subject 6 A good number. For two people to go on holiday together, that is nice. It is a fine number, 2. Two children, that is also good.

Subject 7 Two hands, two feet, two ears, two eyes.

Subject 8 People, things, a pair. Man, woman.

Subject 9 It makes me think of a swan.

Subject 11 I am a football player, and my favourite number is 2. Number 2 means you are a defender.

13

Subject 1 It is an unlucky number. I don't believe in that so much really but …it is prime also and 13, that isn't a good age to be. And then of course it is odd.

Subject 5 For most people it is an unlucky number, but for me it is lucky.

Subject 6 It is an unlucky number of course. I believe in that a little. But on the other hand, if you happen to draw it, it could also mean that you are not unlucky, just because everyone is always talking about it.

Subject 8 Its unlucky. At that age I first went to secondary school.

Subject 9 Unlucky. Adolescents.

Subject 10 An unlucky number. My birth date, so it's lucky for me.

Subject 11 An unlucky number, but that is not true for me.

Subject 15 A terrible number to calculate with. Because it's very, very odd.

Table 4.1 contains some examples of the ideas people come up with when inspecting a number for its meanings. Numbers may stand for certain ages (37 means being settled, 13 is an unhappy adolescent); they may have fixed meaning in certain sports, in this case soccer (number 2 is a defender, number 12 starts on the bench as a substitute); they are either divisible (25) or not (37); they may refer to important social configurations (2 is nice, 1 is lonely), or to currencies (25 is a quarter); and they may have a special look about them (2 is a duck or a swan). And, of course, 13 is unlucky, whether you believe in such things or not. Subjects also use a variety of descriptive terms, such as agile, beautiful, lonely, interesting, good, terrible, and "very, very odd".

In the second part of the interview some follow-up questions were asked: are there any numbers you especially like or dislike, and what are the qualities that make a number more or less pleasant to you? Table 4.3 sums up the statements of all subjects, while Table 4.2 gives the descriptions produced by one highly articulate subject.

Table 4. 2.

One subject's description of preferences and aversions.

Subject 1	I prefer even numbers. Easy numbers, that you can do something with. And 7 that is just… no, that is not a good number to me. It is unfriendly. And 11 also. When you asked me to comment on these two numbers it gave me a bad feeling. Not really terrible of course, but the first thing that occurred to me was ugh, 7 and 11. They seem so closed. If they were people, they would be stiff and surly. But 64, that is a wonderful number. And 100, well that has been a little overdone perhaps, that's a number everyone will always talk about. But 8 is a beautiful number, with a nice look to it too. Mostly, I like numbers below 50, or at least below 100. Up from there numbers have so little that you can think about. Take 172, that is a 1, a 7 and a 2. And 1, that is all right, but 7 is not, and 2 is OK again. So that is difficult to make a unity of. All in all, I prefer numbers below 100, and I like even numbers better than odd numbers.

It is evident that most of the interviewed people had some opinions about numbers. New descriptive and evaluative terms are introduced in Tables 4.2 and 4.3, such as unfriendly, closed, surly, stiff, stupid, nice, round, pleasant and full. Number 8 is a good mark in school, while 1 and 0 mean failure. One subject speaks of the taste a number gives him, one subject refers to a number's sound, other people comment on how numbers look.

It is equally clear, however, that opinions differ: two people prefer even numbers, but one prefers odd. Many people like 8, but one person dislikes it, just as Eberhart did. Number 10 collects favourable mentions only, whereas 1 is exclusively distasteful. The numbers 2, 3, 4, 5, 7, and 13 are agreeable to some but disagreeable to others.

Table 4.3.

An overview of subjects' preferences and aversions

Subject 1	I prefer even numbers. They are easier to handle. Odd numbers are unfriendly and closed (see Table 4.2).
Subject 2	5 is my favourite superstitious number. I like 8 because of its form. And because it is a good mark in school. To me 1 is stupid, and 0 also.
Subject 3	I like numbers that are nice and round. I like 10. It is easy to multiply with, you only have to move the comma. And I dislike fractions.

Subject 4 I like 3, and most numbers that are divisible by 3. I prefer odd to even numbers. Because there is asymmetry in them, which is pleasant to me. I have an aversion of 8. Most numbers with 8's and 2's look ugly.

Subject 5 13 is my favourite, but in general I like even numbers better than odd. And I really hate 1.

Subject 6 I think 8 is a beautiful number. I don't know why. And I like squares. To me 25 and 36 are also beautiful. But 1 gives me a bad taste. Such a very small number.

Subject 7 I like 9, 25 and 23. I don't know why. I dislike 5. Also 13, because it is unlucky.

Subject 8 Numbers that I prefer… 7 is beautiful and 2 is OK. I have no particular dislikes.

Subject 9 8 is a beautiful number, to get as a grade. My lucky numbers are 6 and 3. And 666 has a nice sound to it.

Subject 10 My lucky number is 4. There aren't any numbers I dislike. Of course I recognize some numbers. 3 is the number of god, 5 is the number of Maria, and 6 is the devil's number. Especially 666, as I remember from a movie, The Omen, where it was carved into his head.

Subject 11 I'm attracted to 5. When I must choose a number it is often 5. And 10 is good too, it is so nice and full, and also pleasant to work with. And 2 of course, but that is for football. Though at the moment I'm given shirt number 12. As to aversions, I will never choose 4, and as a child I had a special dislike of 17, because of those stripes that would not fit and just kept hanging in the air.

Subject 12 My own favourite is 13. Because it is unlucky, and I am a little superstitious. But I have no aversions. I don't mind numbers generally.

Subject 13 When I must name my lucky number I always say 8. I don't know why. I once won something with it, but that can't be the reason, because it was my lucky number before I ever did. And I like 64, perhaps because it is the square of 8. I like 256 too. As to aversions, I dislike 3. It seems unfriendly to me. And 24 is a little dull.

Subject 14. There are no numbers I prefer. And I don't dislike any numbers either. Though of course it is convenient when a number ends on 0.

Subject 15 I like beautiful round numbers such as 8 and 10. I mean round to the ear, and to the eye, and also as an idea. And 3 is very ugly indeed. I don't have a big problem with 13, but it doesn't really suit me either.

There are also differences in numbers' frequencies of being mentioned. Top scorer is 8 (7x) followed by 3 (5x), 1, 2, 5, and 13 (4x), 10 (3x), 4, 6, 7, 64, and 666 (2x). Twelve numbers are named once. Most statements are made about the numbers from 0 to 13. An explanation of this phenomenon is given by Subject 1 (see Table 4.2): larger numbers are more difficult to make a unity of. A similar observation was offered by the expert mental calculator who stated that "for him, every number up to a thousand was but one idea, and every number between a thousand and a million was, to his regret, two ideas" (Hunter, 1977, p. 43). Our subject draws the line at fifty, or perhaps hundred. To her most numbers beyond that line

have no individual identity.

The first experiment involves a test of all numbers 1-100 on the Semantic Differential. This test was developed by Charles E. Osgood and other psychologists during the Fifties as an instrument for "the measurement of meaning". It is obvious, they wrote (Osgood, Suci & Tannenbaum, 1957), that the meaning of a concept is not really captured by measures such as Noble's m, which is nothing but an index of its associative value. The words jelly, jewel, and heaven have approximately equal m-scores, but they have quite different meanings. The investigation of meaning must take content into account, but how? A concept's psychological meaning is not equal either to the sum of its publicly available definitions and descriptions. How to capture meaning as it is subjectively experienced? The Semantic Differential was developed and presented as one solution to this problem. It was constructed in a bottom-up fashion. First, hundreds of concepts were rated by groups of American subjects on as many as fifty bi-polar scales. Examples of such scales are good-bad, active-passive, big-small, thick-thin, etc. These data were then factor analyzed. Three major factors were extracted, together with some minor ones. The dimensions represented by those three main factors could be identified by looking at the scales which had the highest loadings on them. After repeated testing the following conclusions could be drawn (Osgood et al., p 72-73). The largest factor to be extracted after simple structure rotation is always an evaluative factor. It explains at least half of the common variance in the data. The scales loading highest on this factor are of the good-bad type. The second factor is a potency factor or, as one of Osgood's co-operators would say, a "football player factor". It is concerned with power and things associated with it, such as size, weight, toughness, and the like. The third dimension is an activity factor, concerned with quickness, excitement, warmth, and agitation. Some examples may illustrate what is meant. A white rose-bud is, by its ratings, viewed as positive, impotent, and passive, whereas a villain is negative, potent, and active. The instrument has been, and still is, used to measure differences in implicitly experienced meaning. For instance, the concept of "federal employee" was shown to have positive connotations for Democrats and negative ones for Republicans (Osgood et al., 1957). Not all scales served equally well to represent a factor. For further measurements, only the most representative scales were selected. A cross-cultural validation study subsequently revealed that the best American scales were not always the best scales elsewhere. In our experiment we made use of the standardized Dutch version of the Semantic Differential, which was developed by Jansen and Smolenaars (1967).

EXPERIMENT I

Method

Stimuli and materials

All numbers 1-100 were used as stimuli in the experiment. Each number was to be rated

on twelve bi-polar scales. These scales used in the experiment were taken from the standardized Dutch version of the Semantic Differential, developed by Jansen and Smolenaars (1967). In this version the three main factors Evaluation, Potency and Activity are represented by four scales each. These were:

1) for Evaluation: pleasant - unpleasant (prettig - naar), sympathetic - unsympathetic (aardig - onaardig), good - bad (goed-slecht), and happy - unhappy (gelukkig -ongelukkig);
2) for Potency: strong - weak (sterk - zwak), large - small (groot - klein), hard - soft (hard - zacht) and heavy - light (zwaar - licht);
3) for Activity: interesting - dull (boeiend - saai), quick - slow (vlug - langzaam), active - passive (actief - passief) and excitable - calm (snel opgewonden - kalm).

Each scale had seven scale points, with one concept printed at the utmost left and one at the utmost right. For each factor two scales were presented in the original direction and two in the opposite one (e.g., good-bad; unhappy-happy). The order of presentation was fixed. The first scale was an evaluation scale, the second a potency scale, the third an activity scale, and so on. The twelve scales were printed on one page. The stimulus-number to be scored was printed at the top of each page.

The stimulus-numbers were randomly divided into ten groups. Ten different booklets were prepared and reproduced.

Subjects

Subjects in the experiment were 450 first year psychology students. Participation was part of a course requirement. Subjects were tested in three groups.

Procedure

The Semantic Differential was included in a large battery of other tests and questionnaires. Five experimenters were present to assist. Subjects sat behind desks, as in an exam. They received one booklet each, and were instructed to rate the ten stimulus-numbers on each of the twelve scales. The whole task had to be completed within ten minutes.

Results

Three principal components were obtained with eigenvalues greater than 1. Factor scores of all stimulus numbers obtained after Varimax rotation are given in Appendix 3. The percentage of the original variance accounted for by these three factors is 87 percent. Their individual contributions are 52, 21, and 14 percent, respectively. As can be seen from Table 4.4, the Evaluation factor is indeed strongly correlated (.91 - .95) with the four scales meant to represent it (pleasant-unpleasant, sympathetic-unsympathetic, good-bad and happy-unhappy), while the loadings of these scales on the other factors are weak. However, four other scales also have their strongest correlations with the evaluative factor. These scales are strong-weak, interesting-dull, quick-slow and active-passive. The second factor, Poten-

cy, has strong and exclusive loadings on two scales: big-small (.85) and heavy-light (.81). Factor 3, representing Activity, has its exclusive representative in one scale only, being the excitable-calm scale. The other variable receiving loadings from this factor, the hard-soft scale (.54), has correlations with Evaluation and Potency which are almost as high (-.40 and .48).

Table 4.4

Factor loadings of the twelve scales on each of the three factors after Varimax rotation.

	Evaluation	***Potency***	***Activity***
Pleasant-unpleasant	.95	-.15	-.11
Sympathetic-unsympathetic	.91	-.17	-.10
Good-bad	.91	-.06	-.22
Happy-unhappy	.91	-.21	-.10
Strong-weak	.80	.34	.17
Big-small	.16	.85	-.23
Hard-soft	-.40	.48	.54
Heavy-light	-.21	.89	-.10
Interesting-dull	.82	.08	.18
Quick-slow	.60	-.43	.57
Active-passive	.74	-.24	.52
Excitable-calm	-.07	-.26	.82

Scales

Mean ratings and standard deviations of the twelve scales are: pleasant 4.38 (.67), sympathetic 4.37 (.50), good 4.40 (.45), happy 4.22 (.40), strong 4.39 (.47), big 4.17 (.65), hard 3.97 (.36), heavy 4.10 (.42), interesting 4.22 (.43), quick 4.03 (.42), active 4.25 (.40), excitable 3.84 (.34).

One scale was chosen to represent each factor. A criterion was that the selected scale should be strongly associated with one factor only. Also, the intercorrelations between the three scales should be as low as possible. For Evaluation the selected scale the good-bad scale, for Potency the heavy-light scale, and for Activity the excitable-calm scale.

A reliability test for these three scales gave alpha reliability values of .81 for the good-bad scale, .93 for the heavy-light scale and .53 for the Calm-Excitable scale. When reliability is measured for the selections of twenty numbers taking the most extreme positions on each scale (those given in Table 4.5), the alphas are .93, .91 and .83 respectively.

Table 4.5.

Numbers taking the most extreme positions on the three bi-polar scales. Rank orderings (first column) are from Good to Bad, from Heavy to Light and from Excitable to Calm.

	Good to bad	*Heavy to light*	*Excitable to calm*
Rank	number	number	number
1	10	87	13
2	100	83	9
3	36	100	81
4	8	67	1
5	24	75	15
6	66	58	11
7	16	92	28
8	4	93	7
9	1	78	33
10	88	65	3
91	41	38	40
92	51	5	72
93	39	15	70
94	17	7	84
95	13	21	6
96	59	3	88
97	29	11	60
98	43	20	36
99	53	4	82
100	67	22	80

Properties of numbers

What does it mean for a number to be good or calm? Or, alternatively, how are such concepts represented within the category of numbers? Some indications may be obtained by an inspection of the sets of numbers taking extreme positions (negative and positive) on three representative scales: good-bad (Factor 1), heavy-light (Factor 2) and Excitable-Calm (Factor 3). The ten numbers standing out on both poles of these three scales are printed in Table 4.5.

An inspection of Table 4.5 suggests that some numerical properties are particularly relevant for the orderings obtained. Good numbers are predominantly even and bad numbers are odd. Tabledness, also, may have some relevance for this scale since none of the ten numbers with the highest bad-scores occur in the multiplication tables, while all good numbers do, five of them being squares. The heavy-light scale is primarily sensitive to size, which was to be expected. The excitable-calm scale seems sensitive both to size and to the odd-even distinction.

These notions were tested by multiple regression. The predictor variables we used were

1. Size, 2. Odd-even, and 3. Tabled versus non-tabled. Multiple correlations and standard regression coefficients beta weights) are given in Table 4.6.

Table 4.6.

Multiple correlations and standard regression coefficients (beta weights) of three predictor variables (size, evenness and tabledness) for ratings obtained on three scales (good-bad, heavy-light and Excitable-Calm).

	R	***Size***	***Odd/even***	***Tabled***
Good	.68	.05	.39**	.47**
Heavy	.74	.66**	-.01	-.14
Excitable	.65	-.34**	-.50 **	.07

** = $p < .001$.

All multiple correlations are significant, $p < .0001$. Table 4.6 indicates that each subjective scale is associated with one or two objective measures. Evenness and tabledness positively influence a number's rating on the good-bad scale. Size is the main objective attribute in determining the heavy-light ratings. Excitable ratings are negatively correlated with size and evenness, indicating that small odd numbers are felt to be more excitable, and large even numbers are felt to be more calm. It is also possible that numbers with a socially shared connotative meaning tend to be small. This is different from stating that small numbers have some general tendency to be rated as excitable. We will return to this matter in the discussion.

Discussion

What is perhaps most remarkable in these results is that the meaning of numbers, as measured by the Semantic Differential, can be described on the same dimensions that have been found to characterize rosebuds, villains and American Senators. evaluation is the largest factor in this experiment, as it is commonly found to be; Potency comes second and Activity third. The "football player" factor, which has to do with size and power, is not much larger here than it is in other experiments. As it stands, the factor has a high though not exclusive correlation with numbers' actual magnitude ($r = .73$). The same is true for the heavy-light scale we selected as its best representative ($r = .72$). Its high reliability can be read as a further indication that subjects used a well known common standard - magnitude - in their heavy-light ratings. Two of the four scales intended to represent the Potency factor fail to do so. These scales also measure something different than size. The correlation between magnitude and hard-soft scores is only .27, and that between magnitude and strong-weak scores is as low as .11.

Evaluation is clearly the most important dimension underlying subjects ratings. It shows many signs of representing some solid and stable component of the meaning of numbers.

All four scales meant to represent the factor have high loadings on it (from .91 to .95, see Table 4.4). The intercorrelations between those four scales are between .87 and .90, which is also high. It suggests that the bipolar distinctions that were offered by these scales were interpreted in very much the same way, and that a good number is also felt to be happy, pleasant and sympathetic. The overlap with four other conceptual distinctions: interesting-dull, strong-weak, active-passive and quick-slow is considerable (intercorrelations with the four evaluation-scales are between .45 and .75). This was unintended, and it has contributed to the factor's strength. The inter-subject consistency of evaluative ratings, as measured by the reliability of the good-bad scale, is quite high (.81 for all hundred numbers and .93 for the top and bottom tens). It must mean that one specific opinion has emerged as the dominant one, and that the people holding that opinion like some numbers well and others not at all.

The best liked numbers are 10 and 100, followed by 36, 8, and 24. These are all "easy numbers that you can do something with", as it was formulated by one of our subjects (cited in the introduction to this chapter, Table 4.2). All top scorers except 1 are even, and most are symmetrical in some additional respect, such as repetition, the second digit being a double of the first or being a square. The complete good-bad scale has a positive correlation of .50 with the frequency scores that were obtained in Chapter 2.

All ten numbers with the highest bad-ratings are odd and relatively large. All but two are prime. Of the two that are not, 39 is a multiple of the "very, very odd" 13, while 51 is a multiple of 17, both of which are also present on the list. Interestingly, the very worst number, 67, is also the one with the lowest m-score (see Chapter 2, Table 2.2). Three other numbers, 53, 43, and 59, also occur on the list of those scoring lowest on Noble's m. In fact m-score and Good ratings of all numbers correlate .46.

The Activity factor is the smallest one and its sole good representative, the excitable-calm scale, is not very reliable, with an alpha of only .53. As we have seen, this measure of inter-subjective consistency in ratings is much better (.83) for the twenty top and bottom numbers. It could well be that the dimension represented by this scale is not equally relevant to all 100 numbers. Of those ranking among the top ten, several have affective connotations which are part of public knowledge. The Winkler Prins Encyclopaedia names 13 as being associated with bad luck, and 3 and 7 as holy as well as magical. Number 9 has some magic too, while 11 stands for madness, sin and intemperance. Four other numbers on the Excitable list are multiples of 3. All except 28 are odd, whereas all numbers with the highest calm ratings are even.

In this experiment all numbers 1-100 were brought to the attention of subjects, who were then asked to rate them. The purpose of the Experiment 2 was to investigate what happens when people are given no such cues, but must search their own minds for numbers to fit some verbal concept. In Chapter 2 we have seen that small numbers were named much more often than large numbers. However, these results were produced by investigating associations between numbers. In the present setting we are concerned with associations

between numbers and certain kinds of affect. Experiment 2 serves two objectives. The first is to obtain an overview of people's choices. Some factual questions (see Method section) will be asked: what is your favourite number, and which number do you hate most. These will be followed by questions that are more tentatively phrased: if you had to name a pleasant number, or an excitable number, which would you choose to fit such a concept? The results of the pilot study reported in the introduction suggest that numbers above 13 have a relatively small chance of being named in such a context. We have also seen that opinions differ: some frequently named numbers were cited as pleasant as often as they were cited as hateful. This being so, one may well wonder if the semantic distinctions as measured in Experiment 1 have any chance of being reproduced in a different setting. The second aim, then, is to address the question whether the three dimensions proposed by Osgood, and found to underlie subjects' ratings in Experiment 1, will stand up to the limiting effects of frequency, in combination with the variance caused by differences in taste.

EXPERIMENT 2

Method

Subjects

Subjects were 472 first year psychology students at the University of Amsterdam. They were tested collectively in three groups.

Materials

A six item questionnaire was prepared. As a first question it was asked if subjects had a favourite number, yes or no. Which of two alternative questions was subsequently presented depended on the answer to question 1. If the answer was yes, the second question invited this subject to write down his or her favourite number, and to give an estimate of the age at which the preference had formed. If the answer to question 1 was no, the subject was invited to produce the number he or she would have chosen for their favourite as a child. All subsequent questions were identical for both groups. Question 3 asked people to write down the number they hated most. Question 4 introduced the notion of pleasant and unpleasant numbers, inviting subjects to produce three examples of both. Question 5 stated that a distinctions sometimes made between interesting and uninteresting numbers. Could they name three of each kind? Question 6 mentioned that subjects participating in a previous study had rated numbers on a scale from excitable to calm. Can you imagine, it said, which numbers are most often associated with these concepts? Name three numbers to match both categories.

Procedure

Subjects used pen and paper to fill out the questionnaire, which was presented as a survey. The introduction stated that we were interested in people's opinions about numbers. They

would be presented with some concepts, such as pleasant and unpleasant, and should produce numbers to match them. Perhaps, it said, you cannot avoid naming some number more than once. In that case you are allowed to do so. "Please take your time to absorb the questions before writing down a response". Subjects were given ten minutes to complete the task.

Results

To question 1, which asked subjects if they had a favourite number, 231 people answered yes and 240 answered no. The subjects who answered yes all named a favourite number. The average (estimated) age at which the preference had been acquired was 10.2 years. Of the subjects who answered no, and were subsequently asked to name the number they would have chosen as a child, twelve failed to do so.

Should each subject have produced all twenty numbers that were asked for, the yes group's total production would have been 4620 and the no group's production 4800. In reality the yes group produced a total of 4310 numbers and the no group 4206. This gives 310 (7 percent) omissions for the yes group, and 594 (12 percent) for the no group. The difference between the percentages of both groups, averaged over the eight categories, is significant, $F(1, 14) = 4.64$, $p < .05$. It means that the yes group had less difficulty with the task as a whole.

Of the 8516 numbers that were produced, 7788 (91.4 percent) are whole numbers between 1 and 100. The percentages are 92 and 91 for the yes and no groups respectively. This difference is not significant ($p > .10$). Of the numbers outside the 1-100 category, 0 is by far the most frequent one, accounting for 299 of the 728 "outsider" responses. The yes and no groups produced this number with about equal frequency (152 and 145 times respectively). Non-natural numbers were produced in 112 cases, which were also equally distributed between the yes and the no group (58 vs 54 cases). Examples of such response numbers are e, p, ÷2, 33.3, 1/1000. The remaining "outsider" responses were numbers above 100. Examples are 1000, 1.000.000, 999, and 666.

Favourite numbers

An overview of subjects' choices is given in Table 4.7. As can be seen, the "top tens" of both groups of subjects are rather similar. Four numbers, 3, 7, 8, and 4, lead the lists of favourite numbers of both groups. However, the rank orderings of these four numbers are different (see Table 4.7). In the yes group number 3 comes first (15.2 percent), followed by 7 (13.0 percent), 8 (10.8 percent) and 4 (9.5 percent). The ranks produced by the no group are: 7 (25.0 percent), 8 (15.8 percent), 3 (10.0 percent) and 4 (9.2 percent). In the yes group these four numbers account for 48.5 percent of all "ffavourite" responses. In the no group they account for 60.0 percent of the "remembered favourite" choices.

Table 4.7.

Present and remembered avourites, with their proportional frequencies. (The ten winning numbers are given for each group, and for all subjects combined. The percentages are based on the total number of participating subjects, including non-responders).

Favourite numbers

	Yes group((n=231)		No group (n = 240)		All (n = 471)	
Rank	Number (%)		Number (%)		Number (%)	
1	3	(15.2)	7	(25.0)	7	(19.1)
2	7	(13.0)	8	(15.8)	8	(13.4)
3	8	(10.8)	3	(10.0)	3	(12.5)
4	4	(9.5)	4	(9.2)	4	(9.3)
5	6	(4.8)	13	(4.2)	5	(4.0)
6	5	(4.8)	2	(4.2)	13	(4.0)
7	13	(3.9)	10	(3.8)	2	(3.8)
8	2	(3.5)	5	(3.3)	6	(3.6)
9	14	(3.5)	6	(2.5)	9	(2.5)
10	9	(3.5)	(1, 9 11)	(1.7 each)*	10	(2.5
Sum of %		72.5		83.1*		74.7

*This column, which has a tie, contains twelve instead of ten highest scoring numbers.

Of the 231 subjects who had a favourite number, 194 could also name a most hated one. Of the 240 subjects without a present favourite 184 could do so. Table 4.8 gives the numbers that were most frequently named as Hated by the two groups. As can be seen from the table there is much overlap between the groups. Nine of the ten most hated numbers occur on both lists. The exceptions are 2 (yes group) and 6 (no group), which, however, do appear somewhere lower in the hierarchy of the group that did not place them in its top ten.

Table 4.8.

Top tens of Hated numbers calculated for the *yes* and *no* groups separately and for all subjects combined.

Hated numbers

	Yes group	No group	All subject
Rank	***Number (%)***	***Number (%)***	***Number (%)***

1	7	(15.2)	3	(11.2)	3	(10.8)
2	3	(10.4)	13	(10.4)	7	(10.2)
3	5	(9.1)	1	(7.5)	13	(9.3)
4	13	(8.2)	7	(5.4)	5	(7.0)
5	0	(6.1)	5	(5.0)	1	(5.7)
6	17	(3.9)	4	(4.6)	0	(5.1)
7	1	(3.9)	0	(4.2)	4	(3.8)
8	2	(3.0)	6	(4.2)	6	(3.2)
9	4	(3.0)	9	(2.9)	17	(3.0)
10	9	(2.6)	2, 9, 11, 17	(2.1 each)*	9	(2.8)
Sum of %		65.4		63.3*		60.9

*This column, which has a tie, contains fourteen instead of ten numbers.

However, there is also considerable overlap between the categories of favourite and hated numbers, 7 and 3 being top scorers in both. Other numbers that rank among the (overall) top tens of both categories are 4, 5, 6, 9, and 13. Only a few numbers seem to be prominent in one category only: 8 is liked by many and hated by few, and 0 and 17 are mentioned frequently in the favourable, and seldom in the unfavourable context. To obtain a clearer picture of relative likes and dislikes we subtracted the "favourite" and "hated" scores of all numbers between 0 and 100. The resulting rank orderings are presented in Table 4.9.

Table 4.9.

Rank orderings obtained after subtraction of "hated" from "favourite" frequency scores.

Positive:	**8** (+55), **7** (+42), **4** (+26), **10** (+10), **12** (+9), **3** (+8), **14** (+7), **24** (+6), **15** (+6), **18** (+6), **2** (+6), **16** (+5) and **25** (+4)
Negative	**13** (-25), **1** (-23), **0** (-22), **5** (-14), **17** (-12), **19** (-7), **37** (-3), **27** (-2) and **67** (-2).

Questions 4-6 asked subjects to name three examples of pleasant, unpleasant, interesting, uninteresting, calm and excitable numbers. Again, a highly limited set of numbers was used by our subjects to represent these categories. The resulting redundancy (or overlap between categories) could be somewhat reduced, however, by limiting the analyses to subjects' first choices for each category. It is evident from Table 4.10, which presents the resulting six "top tens", that even after discounting second and third choices, considerable overlap remains. Most numbers that occur with any frequency at all do so in more than one category, and many are chosen to represent supposedly opposite ones. The main culprits

are again the numbers 3 and 7, which seem to fit any type of description, as they appear on each individual top ten. The numbers 1, 5, and 10, which are present on five of the six lists, come second on this measure of subjective semantic controversy, followed by 2, 4, 6, 8, 9, and 11 (four lists). In fact, only sixteen numbers are used with sufficient frequency to penetrate any of the top ten lists. These are the numbers 0-12, 13, 17, and 100.

A frequency count performed over the combined scores (obtained in the six categories) of all numbers 0-100 confirms these observations. The sixteen highest scoring numbers are, in order of their overall frequency: 3, 7, 2, 1, 8, 4, 5, 13, 9, 6, 10, 0, 11, 100, 17, and 12. Together, these sixteen numbers account for 6378 of all 7977 responses involving any of the numbers 0-100, which is eighty percent. The two top scorers are the numbers 3 and 7, with scores of 697 and 695 respectively. These two numbers alone account for 17.4 percent (8.7 percent each) of all responses to the category questions.

Table 4.10.

Numbers with highest net frequency scores in six subjective semantic categories. Only first named numbers were counted. For each category the ten leading numbers are given.

Pleasant	*Unpleasant*	*Interesting*	*Uninteresting*		*Calm*	*Excitable*
8	13	7	1	2	9	
7	3	3	2	1	13	
3	7	9	3	8	3	
2	1	13	0	4	7	
4	5	8	4	3	1	
9	11	11	5	6	10	
10	17	1	6	10	5	
6	9	2	7	0	11	
5	4	0	10	7	8	
12	6	10	11	5	100	

Table 4.11 gives the rank orderings that result after subtracting, for each number, its scores obtained on opposite categories: pleasant minus unpleasant, interesting minus uninteresting, calm minus excitable. By thus eliminating the most obvious effects of general affective frequency, certain more specific characteristics of numbers may have a better opportunity to present themselves. By this procedure twenty-eight numbers succeed in making an appearance. Twelve of these appear on three lists (which is the highest possible score after subtraction), eleven numbers represent two, while five numbers come out as representatives of one concept only.

Table 4.11.

Top tens of numbers on six lists after subtraction.

	Pleasant minus unpleasant
Pleasant	8, 2, 10, 4, 12, 100, 18, 7, 16, (9&24)
Unpleasant	13, 1, 17, 0, 11, 5, 19, 23, 3, 37
	Interesting minus uninteresting
Interesting	7, 9, 13, 8, 18, 99, 21, 12, 11, (16&69)
Uninteresting	1, 2, 0, 4, 6, 5, 17, 20, 19, (23&27)
	Calm minus excitable
Calm	2, 4, 6, 8, 1, 0, 24, 20, 12, 3
Excitable	13, 9, 99, 11, 69, 7, 100, 17, 5, 88

Net pleasant scores range from 116 (obtained by the response number 8) to 16 (obtained by 9 and 24). Net unpleasant scores range from 111 to 8 (obtained by 13 and 37 respectively). For interesting the range is from 52 to 12, for uninteresting from 90 to 9, for calm from 153 to 13, and for excitable from 94 to 11. In most cases the distance between first and second placed numbers is considerable. The net scores of 8 (pleasant), 13 (unpleasant), 1 (uninteresting), 2 (calm) and - again - 13 (excitable) are almost twice as high as those of their closest competitors.

In Experiment 1 we found that Good numbers are mostly even, whereas bad numbers are predominantly odd. This seems to be true for pleasant and unpleasant numbers as well. All but two of the pleasant numbers printed in Table 4.11 are even. With the exception of 0, all Unpleasant numbers are odd. It is less clear which common characteristics serve to distinguish Interesting from uninteresting numbers. The odd-even distinction seems to lack relevance here. Size may be a factor though, seeing that the three most Uninteresting numbers are 1, 0 and 2. In Experiment 1, the calmest numbers were often even and large, while the excitable numbers tended to be odd and small. Table 4.11 suggests that only the odd-even distinction is reproduced in the present experiment. However, such observations are based on the rank-ordering obtained after eliminating a considerable proportion of the scores of the most frequently named numbers. In the next section the relationships between the objective attributes evenness, size and tabledness (which were also used in Experiment 1) and the subjective qualities measured in the present experiment will be more systematically analyzed. Because 0 cannot be scored on the odd-even variable, and because it was not used as a stimulus in the previous experiments, it was removed from the set of numbers on which the analyses were performed.

Odd-even frequencies

The correlation between the associative frequency scores measured in Chapter 2 and the combined frequencies obtained in the present context is .89. In the continuous association task, which was used as a frequency estimate (see Chapters 2 and 3), the proportions odd and even numbers were 43-: 57, even numbers being in the majority. In the present experiment these proportions have changed in favour of the odd number responses, who now have a 58-:-42 percent majority. Though neither of these differences reaches significance on its own, the interaction between task and numerical attribute does, $F(1, 98) = 16.02$, $p < .001$.

Proportional distributions of odd and even responses within each of the categories are printed in Table 4.12.

Table 4.12.

Proportion distribution of odd and even number responses obtained for eight subjective seman-

tic categories.

	Even	*Odd*	*Individual effect*		*Interaction*
Favourite number	46%	54%			
Hated number	21%	79%	$p < .05$	$p = .06$	
Pleasant	55%	45%			
Unpleasant	21%	79%	$p < .01$	$p < .001$	
Interesting	41%	59%			
Uninteresting	42%	58%			
Calm	64%	36%			
Excitable	32%	68%	$p = .08$	$p < .01$	

The proportions odd and even number responses were compared by one way Anova for each individual category. Opposite categories were also tested for significant interactions by repeated measures Anova. The results are printed in the last two columns of Table 4.12. Significant individual effects of the odd-even distinction were obtained for the categories of hated and unpleasant numbers, in which odd number responses are relatively frequent. A marginal effect was obtained for the excitable category, in which odd number responses also tend to prevail over even ones.

Significant interactions were obtained for the pleasant-unpleasant and the excitable-calm comparisons. As can be seen from Table 4.12, the distribution of even and odd numbers is influenced by these concepts in opposite ways. even numbers are relatively underrepresented in the unpleasant and excitable categories.

Objective and subjective measures

A multiple regression analysis was again performed (compare Table 4.6, Experiment 1) to determine the combined and independent influences of size, tabledness and odd/evenness on the numbers' category scores. The results are printed in Table 4.13. All multiple correlations are significant, $p < .001$. Predictably, in this experiment the size-variable is the largest one for each category, with highly significant negative standard coefficients ranging from -.37 (for the excitable category) to -.53 (for the unpleasant category). There is a positive relationship between the frequency of a number's appearance in the pleasant, interesting and excitable categories and its membership of the multiplication tables. That even numbers tend to be less hateful, less unpleasant and less excitable (see previous section) was confirmed.

Table 4.13.

Multiple correlations and standard regression coefficients obtained by multiple regression. Three

objective numerical attributes were used for the prediction of numbers' frequency of nomination for eight subjective categories.

	R	*Size*	*Tabled*	*Odd/even*
Favourite	.49	–.42**	.13	–.06
Hated	.54	–.46**	.07	–.22*
Pleasant	.59	–.50**	.19*	–.00
Unpleasant	.60	–.53**	.05	–.26**
Interesting	.61	–.47**	.25**	–.16
Uninteresting	.60	–.53**	.13	–.11
Calm	.55	–.46**	.16	–.06
Excitable	.52	–.37**	.22*	–.24*

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

The predictive value of the Semantic Differential

The final question is whether the three semantic factors obtained in Experiment 1 have something extra to contribute to the prediction of the numbers' scores in Experiment 2. To investigate this matter a second series of multiple regression tests was performed, with six predictors, being the objective variables size, tabledness and evenness (see Table 4.13), together with the factor scores of numbers on the factors Evaluation, Potency and Activity. Again, three predictors were found to be statistically relevant to one or more of the eight frequency distributions. This time, however, these were the objective variable size, and the two subjective variables evaluation and activity. Table 4.14 presents the contributions of these three variables, together with the multiple R's.

Table 4.14.

Multiple correlations and standard coefficients obtained for six predictors on eight categories.

	R	*Size*	*Tab*	*Even*	*Evalu.*	*Pot.*	*Activ.*
Favourite	.51	–.50**	.05	–.08	.14	.07	.00
Hated	.57	–.40*	.02	–.12	.07	–.08	.23
Pleasant	.64	–.57**	.03	–.07	.28*	.03	.00
Unpleasant	.64	–.47*	.00	–.14	.08	–.06	.26*
Interesting	.70	–.45**	.08	–.11	.31**	–.06	.23*
Uninteresting	.64	–.51**	–.01	–.11	.25*	–.07	.12
Calm	.60	–.52**	–.02	–.03	.33**	.00	–.03
Excitable	.67	–.33*	.03	–.13	.34**	–.09	.35**

Note. Column headings denote the variables R, size, tabled/ non-tabled, odd/even, *evaluation, potency, and activity.*
* p < .05 ** p < .01

It seems that the factor scores of Evaluation and Activity succeed in capturing the effects of tabledness and evenness, while contributing something extra besides. Potency, on the other hand, has nothing to offer over the contribution of size. Using the scores on the three scales of good-bad, heavy-light and excitable-calm instead of the factor scores does not change this picture in any important way, which confirms that these scales are indeed good representatives of the three dimensions. Table 4.14 shows that a high score on the factor Activity makes a number more unpleasant, more interesting, more excitable and marginally more hated in the second experiment. These connections are fairly straightforward. The contribution of Evaluation, however, is more semantically diffuse. A high Evaluation factorscore not only makes a number more pleasant, which is as might be expected, but also raises its chances of being named as interesting, uninteresting, calm and excitable.

The influence of frequency

The explanation could be that Evaluation, as a factor, is more sensitive than Activity to properties of numbers that also contribute to their general frequency. We investigated this notion by adding the frequency variable (see Chapters 2 and 3) as a seventh predictor in the Multiple Regression analysis. The outcome is presented in Table 4.15. Multiple R's are raised to values between .68 and .91. The contributions of size and the Evaluation factor have become non-significant. Frequency is the highest contributor by far on each measure, with standard coefficients between .71 (favourite scores) and 1.04 (calm scores). Frequency is the sole relevant variable to the prediction of favourite scores. The variable tabledness makes its re-entrance as a significant predictor, this time with negative standard coefficients, on the measures of pleasant (-.16), unpleasant (-.18), uninteresting (-.23), excitable (-.16) and calm (-.25). The third contributor is the Activity factor, with positive beta weights of .24 for unpleasant, .22 for interesting, .34 for excitable and .21 for hated.

Table 4.15.

Multiple correlations and standard coefficients obtained by multiple regression of seven variables on eight categories.

	R	***Freq.***	***Size***	***Tab***	***Even***	***evalu.***	***Pot.***	***Activ.***
Favourite	.68	.71**	–.10	–.10	–.07	–.05	.07	–.01
Hated	.76	.81**	.06	–.15	–.11	–.15	–.07	.21*
Pleasant	.85	.89**	–.07	–.16*	–.05	.05	.03	–.02
Unpleasant	.83	.83**	—.01	–.18*	–.13	–.14	–.06	.24**
Interesting	.88	.85**	.02	–.11	–.10	.08	–.06	.22**

Uninteresting	.91	1.02**	.06	–.23**	–.10	–.02	–.07	.10
Calm	.89	1.04**	.07	–.25**	–.01	.05	.00	–.05
Excitable	.86	.84**	.14	–.16*	–.12	.11	–.08	.34**

Note. Column headings denote the variables R, size, tabled/ non-tabled, odd/even, *evaluation, potency, and activity*.
* p < .05 ** p < .01

Removing tabledness from the model (see Table 4.16), which leaves us with frequency and the Activity factor as sole predictors, only slightly lowers multiple R's, from an average of .83 tot an average of .81. It reduces the standard coefficients of frequency, which had been artificially boosted by making the influence of tabledness a negative one. The contribution of the Activity factor is raised, and reaches significant values for the categories of hated, unpleasant, interesting, uninteresting, and excitable numbers.

Table 4.16.
Multiple correlations and standard coefficients of frequency and *activity* on eight subjective categories.

	R	***Frequency***	***Activity***
Favourite	.67	.65**	.08
Hated	.71	.63**	.26**
Pleasant	.84	.84**	.05
Unpleasant	.77	.67**	.32**
Interesting	.88	.81**	.27**
Uninteresting	.87	.85**	.16**
Calm	.87	.87**	–.01
Excitable	.84	.71**	.39**

* p < .05 ** p < .01

General discussion

Some of the results may speak for themselves. That around fifty percent of our subjects answered yes to the question if they had a favourite number at present, is a fact which cannot be disputed, and needs not further be commented on. It is also a fact that of all numbers entered into the competition by these subjects 3 scores highest, followed by 7. However, such a fact could be misleading when offered in isolation. For one thing, most subjects have a different favourite. For another, many have a pronounced dislike of 3, while 7, also, is often mentioned as a hated number. What may be concluded, then, from the observations and analyses presented in the results sections?

The simplest conclusion is perhaps that people use a relatively limited set of numbers to

represent their affect. In that respect the results also speak clearly. The frequency distribution obtained in the present study correlates .89 with the one produced in the context of number association, which is a very different task. It shows pronounced similarities also to the distribution obtained by Dietz who, more than sixty years ago, invited people to name the first number that occurred to them. Dietz found that most people named a number under twenty (Dietz, 1933; see also Kubovsky & Psotka, 1976). Such results suggest that the numbers we know well enough to use for some new and under-specified purpose are very few indeed and that they are all small, with the notable exception of 100. Even with complete freedom of choice, we have not much to choose from. The line cannot be definitely drawn at any magnitude however, since there is always a chance that some larger number will be chosen by some person to represent some category. That odd numbers were more often named in the present affective context than they were in the numbers association task can also be concluded with some confidence. However, in the experiment by Dietz, where likes and dislikes did not come in, 7 and 3 were also named most frequently

The observation that even numbers are found to be less unpleasant than odd ones can only be offered with several provisions. Obviously, it should be borne in mind that such a statement generalizes and over subjects, and over numbers. This is, of course, a normal procedure in many experiments. In Chapters 2 and 3 we did exactly the same thing: scores of numbers were calculated by averaging over subjects, after which the means of the categories of tabled and non-tabled numbers were compared. The conclusion that people have less of a problem associating to tabled than to non-tabled numbers generalized over different subjects and different numbers also, and the main purpose of a statistical test is to determine whether such generalizing conclusions may be validly drawn. However, there is an important difference between the two experiments. In the association experiments no subject ever denied that 9 is related to 3, or 20 to 10, though many people failed to make the connection. In Experiment 2 of the present study, several subjects chose 8 and 10 to represent the unpleasant category, thus denying that they are pleasant. For many numbers there is a close competition between subjects who liked them and subjects who explicitly stated that they did not. Such differences in taste are an essential part of the picture produced by the two experiments in this chapter. Numbers, it seems, bear a resemblance to such categories as colours, flowers, books, foods and sports. There are differences in general popularity between the members of these categories, as well as differences in personal taste.

Of course, such differences in taste could also be observed in the ratings of numbers on the scales of the Semantic Differential. Number 10, for example, the one with the highest average "good" score, was nonetheless rated as very "bad" by one subject, and as somewhat more bad than good by another. In Experiment 1 people only made relative judgements, while in Experiment 2 they were forced to state some definite commitments. One may wonder if the first type of rating is not in fact better suited to measure feelings as they really are. Many judgements obtained in Experiment 1 were of the in-between type and few

subjects used the extremes of the scales in any case at all. Probably many people's feelings about the numbers they mentioned are not in fact as pronounced as the words we used suggest.

Nonetheless, most subjects evidently have some feelings about at least some numbers, and the finding that the scores of numbers on the factors Evaluation and Activity of the Semantic Differential have predictive value for the category choices obtained in Experiment 2 is an interesting one. While the factor Evaluation seems to tap characteristics of numbers which have to do with symmetry, and ease of handling (which, presumably, also contributes to their general frequency), the Activity dimension taps connotations of oddness or weirdness that some numbers seem to have. Cultural connotations of bad luck, holiness, magic and madness may be based on the same intrinsic qualities of numbers such as 13, 11, 7, and 3 that determine their higher than average scores on the factor Activity..

Main findings

- When measured on the Semantic Differential, the connotative meanings of numbers fall into three dimensions, represented by the factors Evaluation, Potency and Activity. These same three dimensions have been found by Osgood to underlie the connotative meaning of other verbal concepts. The Evaluation dimension contrasts round and frequent numbers, such as 10, 100, and 36, with infrequent large primes, such as 67, 53, and 43. Potency primarily contrasts between larger and smaller numbers. The Activity dimension selects numbers with culturally established affective meanings, such as 13, 7, and 11, contrasting these with larger, even, numbers such as 80 and 82.
- When affective meaning is measured by a category production task, asking for favourite and hated numbers, as well as for examples of other subjective categories, the chosen numbers are predominantly small ones. In fact, the competition is virtually limited to numbers under twenty. The numbers 3 and 7 are named most frequently of all.
- Associative frequency, which also favours small numbers, is the strongest predictor for each of the affective categories in the production task. However, the Activity dimension of the Semantic Differential also influences people's choices. As a predictor variable, it makes an independent and significant contribution to several categories, most notably to those of excitable and unpleasant numbers. A combination of two experimental measures, associative frequency and Activity factor scores, gives a better description of the category data than a combination of the three objective measures size, tabledness and evenness.

1 Some of the results reported in Chapter 4 have also been published elsewhere. See Milikowski (1992); Milikowski and Elshout (1991).