COHEN AND HANSEL

RISK AND GAMBLING

Dr. C. H. BAKER

Chas Baker
H.F.R. INC.
3040 STATE ST
Santa Barbara
Calif.

RISK and GAMBLING

RISK and GAMBLING

The Study of Subjective Probability

by

JOHN COHEN
Professor of Psychology
University of Manchester

AND

MARK HANSEL
Lecturer in Psychology
University of Manchester

LONGMANS, GREEN AND CO
LONDON · NEW YORK · TORONTO

LONGMANS, GREEN AND CO LTD
6 & 7 CLIFFORD STREET LONDON W I
BOSTON HOUSE STRAND STREET CAPE TOWN
531 LITTLE COLLINS STREET MELBOURNE

LONGMANS, GREEN AND CO INC
55 FIFTH AVENUE NEW YORK 3

LONGMANS, GREEN AND CO
20 CRANFIELD ROAD TORONTO 16

ORIENT LONGMANS LTD
CALCUTTA BOMBAY MADRAS
DELHI VIJAYAWADA DACCA

First Published 1956

PRINTED IN GREAT BRITAIN BY
SPOTTISWOODE, BALLANTYNE AND CO LTD
LONDON AND COLCHESTER

TO

E. AND G.

Speme, diss' io, è une attender certo
della gloria futura, il qual produce
grazia divina e precedente merto.

Dante, *Paradiso*, Canto XXV.

PREFACE

This book is intended for all those who have occasion to fill in football pool coupons or cross a road. We hope that besides being of value to gamblers, punters and professional risk-takers, it may appeal also to those interested in themselves or their fellows. As we have tried to follow the stages of intellectual growth through the school years, we hope that our conclusions may be of some assistance to teachers in understanding the complexity of the child's thinking in relatively simple situations.

In describing our experiments we have attempted to divest them of that technical jargon and statistical mystification which often make utterances on psychological topics unintelligible. The reader will perhaps forgive us if we ourselves have occasionally been unable to resist the temptation to introduce a little non-technical jargon of our own.

With the notable exception of the important work by Professors J. Piaget and Bärbel Inhelder (*La Genèse de l'Idée de Hasard chez l'Enfant*, Paris: Presses Universitaires de France, 1951), a literature on the subject-matter of this book hardly exists. We have therefore limited the text to a description of our studies at Manchester. The reader seeking further information about the experiments will find some of them fully reported in the relevant journals.

We are grateful to the Education Authorities at Manchester, Stockport and Bolton and to the teaching staffs of their schools for permission to conduct our experiments. We are equally indebted to Dr. Eric James, High Master of the Manchester Grammar School, Miss E. M. Bain, Principal of the Withington High School for Girls, and Miss A. M. Bozman, Principal of the Manchester High School for Girls.

For permission to quote from their publications our acknowledgments are due to a number of authors and publishers: Prof. D. W. Harding and Messrs. Hutchinsons Ltd., for allowing us to quote a passage from *Social Psychology and Individual Values*; Dr. W. Grey Walter and Messrs. Duckworth Ltd.; Messrs. Routledge & Kegan Paul; Messrs. W. W. Norton Ltd.; and Dr. Boris Semeonoff, Acting Editor of the *British Journal of Psychology* (General Section).

We are thankful as well to our Research Assistants Mr. E. J. Dearnaley, for help in the computations and for collaborating in the study of pedestrian behaviour, and Mr. J. D. Sylvester, for collaborating in the experiment on beam-jumping.

Valuable aid in the enormous task of tabulation and preliminary analysis was given by Mr. D. B. Walker and Mrs. M. Butterworth. Last but not least, we express our gratitude to Miss Nora Williamson for her painstaking and devoted secretarial help and care in the difficult task of typing the manuscript.

J. C.
M. H.

Manchester, 1955

CONTENTS

Chapter One

INTRODUCTION

In 1654, almost exactly three hundred years ago, the Chevalier de Méré, a mathematically inclined gamester, found himself in difficulties over the problem of dividing the gains in games of dice. So he sought the advice of the greatest mathematician of his time, Pascal; and started a train of thought which has led, during the last three centuries, to the development of the immensely important theory of probability. We may contrast this mathematical probability with the private or psychological world of subjective probability, based on partial or imperfect knowledge, and embodied in the daily perceptions, choices, predictions and decisions of the individual. In spite of the fact that we constantly have to choose, estimate, predict, judge, decide and act on partial or imperfect knowledge, the world of private uncertainty remains largely unexplored.

It is sometimes said that when we use ideas of probability in everyday life we are simply showing how little we know. If, for instance, I say "I am uncertain", I am said to be merely reflecting the imperfect information at my disposal. Laplace declared that "chance is but the expression of man's ignorance". In this way the mysterious world of subjective probability might be dismissed as lacking in interest for science. The demonstrable consistency in our use of subjective probability by itself makes this assumption untenable. Apart from this, we cannot ignore the pervasive influence of subjective probability in all our affairs. There is an element of private uncertainty in most of the things we have to decide about in everyday life. In the minor and major situations which we constantly have to face, we have to choose, predict or act on the basis of partial knowledge. A judge and jury weighing the prisoner's innocence or guilt, a scientist assessing his evidence, a politician predicting social trends, and a gambler wondering what stakes to play, all endeavour to draw conclusions out of the uncertainties which they feel about the situation.

A scientist may be forced to rely on subjective probability because no objective probability is available to guide him. During the second half of the nineteenth century there was a widespread passion for squaring the circle. It was only in 1885 that this was *proved* to be impossible. But before then mathematicians thought it most unlikely, and the French Académie des Sciences year after year used to reject the treatises on this subject submitted by sanguine numerologists or pseudomaths. Poincaré[1] remarks that the Academicians could not have *proved* that they were right. If challenged they would have had to fall back on subjective probability with some such argument as: "We have compared the *probability* that an unknown scientist should have found out what has been vainly sought for so long, with the *probability* that there is one madman the more on the earth, and the latter has appeared to us the greater" (our italics). The word 'probability' is used here in a psychological not mathematical sense.

[1] Poincaré, H., *Science and Hypothesis*, p. 192. London: Walter Scott, 1905.

What delightful uncertainty excites the maiden when she tears petals from a flower and hopefully repeats "He loves me, he loves me not"! Many a popular rhyme [1] embodies the same idea.

The origin of the word 'ponder' suggests that the very process of thinking itself may be an unconscious weighing of uncertain considerations. An entire linguistic mood, the subjunctive, expresses states of uncertainty. Indeed the very fabric of society is made possible by the fact that we deal with other people on the basis of what we expect of them and these expectations always have some element of uncertainty about them however much they may be supported by our experience. The expectation attached to a person's future behaviour is perhaps the most familiar of the uses of subjective probability in everyday life. So seriously does Somerset Maugham [2] treat this expectation that he accuses Dostoevsky of outraging probability in his characterization of Ivan in *The Brothers Karamazov*. Ivan, says Somerset Maugham, is very intelligent, prudent, ambitious, and persistent in his aims. So his vacillation, when hearing of his father's murder is unaccountable; Dostoevsky failed to avoid the *improbabilities—improbabilities* of character, *improbabilities* of incident (our italics).

The idea which led to the present studies began a few years

[1] Tinker, tailor, soldier, sailor,
Richman, poorman, beggarman, thief
This year, next year, sometime, never
Silk, satin, muslin, rags
House, mansion, pigsty, barn
Coach, carriage, wheelbarrow, cart
One, twins, triplets, quadruplets.

[2] Maugham, Somerset, *Ten Novels and Their Authors.* London: Heinemann, 1954.

ago when one of us spent some time in another country. It struck him while he was there that the way people drove their cars about had something in common with the way the foreign policy of the country was being conducted. No driver was satisfied unless the nose of his car was practically touching the tail of the car in front, regardless of the speed at which they were both travelling. What we should regard as the riskiest situations were entered into without a second thought. In foreign affairs as well, those responsible appeared to reduce safety margins to the barest minimum. The policy seemed to be to venture as far as possible without worrying whether any scope for political manœuvre would be left if the policy failed or if the country to which the policy was directed suddenly changed its mind.

It is not important whether these particular impressions were or were not correct. We refer to them because they suggested the possibility that the most diverse of our activities may share some common character which we may refer to as a risk-taking character and which may vary from culture to culture as well as from person to person. Any activity into which we enter with some uncertainty as to the outcome involves an element of risk. In this sense few aspects of life are free from risk-taking. For example a man has to catch a train at 9.0 a.m. What time will he leave his house in order to catch the train? Will he arrive at the station the night before and wait on the platform the whole night? Or will he rush into the station at the last moment just in time to jump on the train as it leaves the platform? Risk-taking may be regarded as subjective probability in action. If we want a definition of risk-taking, we can hardly do better than accept the following one given by a six-year-old boy in reply to the question: "What does 'to take a risk' mean?" He said: "To take a risk is to do something that you are not sure that you are able to do. You are only trying to do it."

Risk-taking as thus defined means doing something whilst tacitly estimating, on incomplete data, your capacity to do it. 'You are not sure' you are able to do something which you are about to attempt. In addition to this type of risk-taking, we should perhaps distinguish another type—that which may involve danger and not simply uncertainty.

The child used the expression 'not sure'. What did he mean by that? When asked, he replied: "I think, I'm sure, I'm certain, I'm positive". He uttered each of these four assertions very differently, as though moving with increasing conviction from the first to the fourth. The tone of voice in which he said "I think" was uncertain and hesitant. The way he said "I'm sure" had firmness about it. "I'm certain" was appreciably more definite still. And when he said "I'm positive", he did so with an air of finality. There appeared to be a continuous scale in the child's mind from uncertainty at one extreme to conviction at the other.

This reply of the six-year-old boy has led us to the study of the language of uncertainty in childhood. A variety of statements are presented involving such categories as quantity, frequency, or duration. Each child assigns a quantitative value to the terms or phrases in a series ranging, say, from 'few' to 'many' or from 'very infrequent' to 'very frequent', or from 'a very short period of time' to 'a very long period of time'. From this data we shall later attempt to trace the changes in quantifiable meanings assigned by children at different ages.

Certain writers on mathematical probability have indeed spoken of 'subjective probability'. By this they have meant that if we know *nothing at all* about the probability of two outcomes, then both are equally possible. Equal ignorance means equal possibility. Personal expectation of an outcome constitutes its 'subjective probability'. We assert that Marmaduke Smith will or will not pass an examination. If we know absolutely nothing else about Marmaduke Smith other than that he is going to take an examination, and if all we know about examinations is that they can be passed or failed, then, in the view of these writers, it would be reasonable to consider Marmaduke Smith's chances of passing and failing as equal. Anyone who knew something about his aptitudes or more than we do about examinations could make another assertion about the probabilities based on *his* personal expectations. In our present usage of the expression 'subjective probability' we mean something different. We mean the expression of our states of uncertainty. In our study of subjective probability we investigate the psychological tendencies by which it is determined. Mathematicians have sometimes taken it for granted that 'psychological laws are much too complicated to fit the scheme of mathematical probabilities'.[1] Our task will be to show that an experimental attack on the problem will prove this view unjustified.

More particularly, our aim in the experimental study of subjective probability and risk-taking is, first, to discover the principles underlying the way we *actually* choose, estimate, predict, judge or take risks and, second, to trace the characteristic changes in these activities during the period of develop-

[1] Reichenbach, H., *The Theory of Probability*, p. 38, transl. by E. H. Hutten and M. Reichenbach, Berkeley: University of California Press, 1949.

ment. We shall see whether judgments based on partial information follow their own psychological rules in more or less consistent fashion and how these rules compare with rules of mathematical probability. We shall also have to consider the way the child assimilates statistical rules in so far as they are repeatedly verified in his experience from the earliest days of life. And we shall be concerned not only with the decision or judgment as such but also with the intensity of conviction with which it is held.

Many of our experiments were carried out in schools and we had to face the possibility that what we told the children, acting as experimental subjects, might not be believed by some of them. They might suspect some deception on our part. We therefore took special care against such suspicions by asking a member of the group to take part in conducting the experiment. Or we invited the teacher to participate so that he could confirm what we told the class was, in fact, true. If, for instance, we had in front of us an urn and beakers containing beads, the teacher would tell the class that what we said about the urn and the beads was correct. After the experiment children could come and look into the urn and beakers or examine other materials and check our statements to their own satisfaction.

A distinction between the children and adult subjects is worth mentioning. In a number of experiments we told our subjects that each bead in the urn was either blue or yellow but no other colour. Then they had to guess how many of the beads in the urn were blue and how many were yellow. In giving their estimates not a single child ever put down any colour other than blue or yellow. But the adults, teachers and prospective teachers, in many cases stated that in addition to so many blue and so many yellow beads, there were red or pink beads in the urn.

A feature of special interest arises in connection with these experiments. Often we gave the subjects no information about the contents of the urn or about the proportions of beads of different colours. They generally assumed that there were equal numbers of blue and yellow (or red and green) beads in the urn. We must be careful to note that this assumption of equal numbers or equal proportions only holds where no information at all is available about proportions and when past experience is no guide. In other situations, subjects may bring in assumptions based on their past experience or drawn from some other source which enable them to interpret the outcomes to their own satisfaction. This tendency may sometimes have rather far-reaching effects. Towards the end of 1941, before the Japanese entered the war, journalists, cartoonists and political publicists in the West displayed considerable uncertainty about the intentions of the Japanese Government. This uncertainty was represented as uncertainty and doubt on the part of the Japanese rulers. The impression was given that the Japanese themselves did not know whether they intended or did not intend to enter the war. The uncertainty of the writers was projected unwittingly on to the Japanese. In so far as it was generally believed in Western countries that this interpretation of Japanese policy was correct and that the Japanese really did not know or had not made up their minds, to that extent there might have been a lack of preparation against Japanese intervention.

In concluding this introduction we might perhaps mention that it is possible to arrive at some rather strange conclusions if we over-value subjective probability. Over-valuation of subjective probability is exemplified by the zeal to square the circle which moved many of our grandfathers. De Morgan compared such zeal to the gallantry of the knight of romance who could not pass by any castle occupied by a giant or magician. During a lecture on the subject of quadrature a member of his audience was heard to mutter: "Only prove to me that it is impossible, and I will set about it this very evening!" De Morgan also tells us about a tract which appeared a hundred years ago, called *The two estates Or both worlds mathematically considered*. The author argued that we can afford to ignore all the happiness that this world can offer if we compare it with the present value of eternal happiness in the future estate. He believed that, like Newton, he could make dazzling discoveries by neglecting infinitely small quantities. His argument ran: let us represent all that this world can give by the symbol a, the present value of the future estate as x, and it will follow that since x is infinitely greater than a, $x + a \equiv x$; a may therefore be safely ignored. This remarkable conclusion did not shake his belief either in the use of mathematics or in eternal happiness.

Chapter Two

PREDICTING THE UNKNOWN

It is a most intriguing thought that many people find it exceedingly hard to grasp that the outcome of an event can be independent of what has happened in similar circumstances in the past. They feel forced to believe that after a run of successes, a failure is more likely to occur than success; and the other way round: after a run of failures, they feel that success is more likely to occur than failure. If a coin is tossed and repeatedly turns up head, most people think that the chance of a further head gets less with each additional toss of the coin. Even those who are familiar with the theory of statistical independence often involuntarily share this belief. The belief is like an optical illusion. Even though we know it to be false we cannot dispel it from our minds.

There seems to be an inability to regard an independent event as separate and detached from a series of similar events in which it occurs. The difficulty remains although it is quite impossible to see how the outcome of the event in question could conceivably be affected by the outcomes of earlier events in the series. The entire series of events is treated as if it were one single continuous event, and the idea of *independence* is displaced by that of *dependence*. Events with outcomes which are independent of previous outcomes are not distinguished from events with outcomes which are more or less dependent on previous outcomes. Both kinds of event are treated as dependent. The great unlikelihood of getting, say, a run of ten heads is attached to the possibility of getting a head on the tenth toss, after a run of nine heads. Contrariwise, the extreme unlikelihood of *not* getting a single head in ten consecutive tosses is attached to the tenth toss after a run of nine tails. As a rule people do not regard the tenth toss as an independent event. So they do not assign equal odds to each of the two possible outcomes.

Philosophers sometimes point to this curious belief as one more proof of human stupidity. But even a philosopher would hardly be content to leave it at that. For the tendency to regard a series of separate independent events as a single continuous event is deep-seated. And it appears to characterize our judgments about a wide variety of different situations in which we assign to one member of a series of events odds properly attaching to the entire series.

Now a 'mathematical man' would not behave in this way, any more than a computing machine would. A 'mathematical man' would base his ideal predictions on formal reasoning and the known properties of things. In the experiments which follow we shall be concerned not with a 'mathematical man' but with the guesses and predictions actually made by the ordinary person. If such predictions differ from formal ones, we shall ask how and why they differ.

There are countless events in everyday life that can have one of several possible outcomes. A race will be won by one out of many horses, the victory in an election for Parliament will be gained by one out of several candidates, a man will choose one of many girls as his bride. There is a particular class of these events which can happen in one of only *two* ways. We shall call these *binary* events. For example, it may rain or not rain on a given day. A child may be born male or female. A student may pass or fail an examination. A story submitted to the editor of a journal may be accepted or rejected. A lady may decide to buy or not buy a certain hat. Such events may occur repeatedly. There might be a succession of similar outcomes—a series of male births or a succession of varying outcomes—male and female births in irregular sequence.

There are two kinds of binary events which we must distinguish at this point. In one kind of binary event we expect the outcome to resemble previous outcomes. We expect a horse that has won many races and lost few to be more likely to win the next race than a horse with a record of many failures and few successes, other things being equal. We believe, and with justification, that our acquaintance with the horse's record of past achievement tells us something about its capacity for winning a race. In the other kind of binary event, such as coin-tossing or die-throwing, the outcome is, in fact, *not* related to previous outcomes. One might imagine that this is what people actually tend to believe; and the universal practice of coin-tossing to decide an issue seems to show that we do not think that knowing the outcome of previous tosses will help us to guess whether the coin will fall head or tail at the next throw. But as we shall soon see, our beliefs and predictions about coin-tossing and similar events are much more intricate than this common practice would lead us to suppose. We shall begin by examining the way children at different ages predict the outcome of the two kinds of binary event when they are given varying amounts of information about previous outcomes.

Prediction of an Independent Binary Event on the Basis of Previous Outcomes

In these first experiments we are concerned with the way children predict the outcome of an independent event that can happen in one of two ways: a bead drawn from a given beaker can be either blue or yellow. Three beads are drawn singly from a beaker and shown to the children, who then have to guess whether the fourth bead will be blue or yellow.

The experiment was designed in such a way that in each arrangement of three beads drawn from a beaker there were always two of one colour and one of the other in the six possible arrangements:

1. Yellow Blue Blue
2. Blue Yellow Yellow
3. Yellow Blue Yellow
4. Blue Yellow Blue
5. Yellow Yellow Blue
6. Blue Blue Yellow

Each of the six arrangements was repeated twice, the whole series of 12 being given in random order, as indicated below:

Y B B
Y B Y
B Y B
B Y Y
B B Y
B Y B
Y Y B
B Y Y
B B Y
Y B B
Y Y B
Y B Y

The beads were drawn from each beaker in the order reading from left to right.

The actual instructions given to the children were as follows. The experimenter says:

> I have 12 beakers in a row. Each beaker contains a lot of beads, which may be blue or yellow, but the proportions of blue and yellow beads vary from beaker to beaker. I am going to take out four beads, one at a time, from each beaker. I will show you the first three beads as they are drawn and then you will guess the colour of the fourth.

The colour of the fourth bead was not revealed.

It will be convenient to call the colour represented by the *two* beads the *preponderant* colour, and the colour represented by the *single* bead the *non-preponderant* colour.

We may describe the aim of this experiment more generally in these words. Children have to guess the outcome of a binary event after seeing a sequence of previous outcomes. The sequences are varied so as to find out the possible effect of particular sequences on the guesses.

The reader will notice that if we ignore the colour of the beads as such and merely think of each bead drawn as being of the preponderant or non-preponderant colour, then the six possible arrangements reduces to three.

Y B B or B Y Y
Y B Y or B Y B
Y Y B or B B Y

If we symbolize a preponderant outcome as 1 and a non-preponderant as 0, the three possible arrangements when colour as such is disregarded are:

0 1 1
1 0 1
1 1 0

Each of the three different kinds of arrangement may have either blue or yellow as the preponderant colour. In analysing the guesses of the colour of the fourth bead we can combine the arrangements in the manner shown and so remove any possible effect due to a preference for one of the colours.

Table 1.

Proportion of preponderant and non-preponderant colours predicted for the fourth bead

Number in sample = 93. Age: 10+

Colour of fourth bead	YBB or BYY 011	BYB or YBY 101	BBY or YYB 110	Total	
				No.	Percentage
Preponderant	0·27	0·30	0·39	362	0·32
Non-preponderant	0·73	0·70	0·61	754	0·68
Total	1·00	1·00	1·00	1116	1·00

As stated, each child made 12 predictions, two for each particular arrangement of beads by colour. As there are now only three possible arrangements if we disregard colour, there will be four predictions made by each child. The sample consisted of 93 children (aged 10+) so there are 372 predictions in all.

We are now ready to study these predictions in order to attempt to answer the two questions:

(i) whether there is a tendency to choose blue more frequently than yellow for the fourth bead when there have been two yellow beads in the first three; and *vice versa* whether there is a tendency to choose yellow more frequently than blue for the fourth bead when there have been two blue beads in the first three, or

(ii) is the prediction for the fourth bead affected by the order in which the two blue (or two yellow) beads have appeared? In other words, does the prediction vary with the three possible arrangements 0 1 1; 1 0 1; 1 1 0?

The answers to both these questions can be obtained by inspecting the figures in Table 1.

It will be seen from the above table (Table 1), by comparing the first row with the second, that there is a marked tendency to predict the non-preponderant colour for the fourth bead. If we consider all three kinds of arrangement combined, as in the last column of Table 1, two-thirds of the predictions are for the non-preponderant colour, which may be either blue or yellow, as compared with one-third of the predictions for the preponderant colour.

Clearly the colour of the fourth bead is not treated as an independent outcome uninfluenced by the colour of the three previous beads drawn. Otherwise there would be no preference for the non-preponderant colour for the fourth bead: there would be no difference between the number of preponderant and non-preponderant predictions respectively. The colour of the fourth bead seems to be guessed on the basis of a sort of principle of compensation or balance, because it tends to equalize the two possible outcomes.

In reply to our second question above we note in Table 1 that the order of appearance of the two colours in the first three beads affects the colour chosen for the fourth bead. The longer it is since the non-preponderant colour appeared, the more often it is chosen for the fourth bead, an effect similar to the one to which we refer in the opening paragraph of this chapter.

We may anticipate a difficulty which may occur to the reader at this point. He may argue that the values for the prediction of the fourth bead shown in Table 1 are averages of 12 separate predictions and that later predictions are affected by earlier ones. In reply it should be made clear that precisely the same kind of preference for the non-preponderant colour was found on the very first occasion, when the order of presentation was 1 0 0. In this instance 0·73 or 73 per cent. of the predictions favoured the non-preponderant and 0·27 or 27 per cent. favoured the preponderant.

We can think of three explanations for the tendency to choose the non-preponderant colour for the fourth bead. These explanations are not mutually exclusive. First, since the children are not told what proportions of blue and yellow beads are in the urn, they might assume, tacitly or otherwise, that the proportions are equal.

From other experiments of ours we know that when the proportions of blue and yellow beads in the urn are stated to be unknown, the children tend to assume that the proportions are in fact equal. Now if, indeed, they made such an assumption in the present case, it would not be contradicted by the fact that two blue and one yellow bead were drawn from the urn. For if only three beads are drawn a ratio of two of one colour to one of the other is the nearest that it is possible to obtain to equal ratios of blue and yellow. The choice of yellow for the fourth bead is thus perfectly consistent with the supposed initial assumption and, moreover, makes it possible for the children to continue to maintain the same assumption. Second, on this same assumption, if, for instance, only one yellow is drawn as compared with two blue beads, the initial store of beads has been unevenly depleted. It would seem to the children that more yellow than blue beads remain in the beaker and therefore it is more likely that the fourth bead will be yellow. Third, as we shall see in experiments to be described later, children dislike unbalanced or asymmetrical arrangements in the form of unequal numbers of blue and yellow beads in any given arrangement. The choice of yellow in the fourth choice when there has been one yellow in the previous three beads (and *vice versa*, of blue for the fourth choice when there has previously been one blue) enables the children to remove this asymmetry.

Exactly the same inclination to prefer the non-preponderant outcome in the fourth event appears if we replace the drawing of beads from urns by real events. For example, we present the following problem:

> Oxford won the boat race this year and last year. Cambridge won two years ago. Which crew do you think will win next year?

The high proportion of children (62 per cent) who believe that Cambridge (the non-preponderant) will win is almost as great as the proportion favouring yellow in the fourth choice, when two blue and one yellow have already been drawn.

Degree of Independence

We have now reached a point when we can introduce a factor which will enhance still more the tendency to predict as the outcome that which has previously happened less often. In the experiments so far dealt with, the earlier outcomes were observed by the children or described to them verbally before they made their predictions. Suppose we give them a visual illustration of previous outcomes, how would this affect their guesses? Let us take as examples those that come within the common experience of children and let us include events ranging from dependent ones, such as sending a present to a person, at one extreme, to independent events, such as coin-tossing, at the other. For instance we say to the class of children:

> A boy has two uncles, Uncle George and Uncle John. In 1951 he received a present from Uncle John but not from Uncle George. In 1952 he received a present from Uncle George but not from Uncle John. In 1953 he also received a present from Uncle George but not from Uncle John. In 1954 he has received a present and on the label there are the words "From Uncle". Which uncle do you think sent the present?

Comparable situations are presented where the prediction refers to winning a toss, to the sex of a baby born in a given town, and to the state of the weather (fine or not fine) on a given day.

In each case the previous outcomes are presented visually thus:

	Presents sent by:	
	Uncle George	Uncle John
1951	—	√
1952	√	—
1953	√	—
1954?		

When we examine the children's predictions made in these *visually illustrated* situations, we find a still more marked tendency to favour the non-preponderant outcome in the fourth choice. Previously, when the situation was not visually illustrated, about two-thirds of the children favoured the non-preponderant, now about 90 per cent. favour it. This heightened preference seems to express a wish to complete a pattern. At the same time it strengthens the tendency to equality or symmetry in the two possible outcomes.

The visual pattern of previous outcomes which is presented also seems to increase the tendency to alternate in guessing from one outcome to the other. For example, if we compare the predictions made when the three following patterns are presented, the tendency to favour Uncle John is most marked in the one on the extreme right (Pattern C).

	Pattern A.		*Pattern B.*		*Pattern C.*	
	Uncle George	Uncle John	Uncle George	Uncle John	Uncle George	Uncle John
1951	—	√	√	—	√	—
1952	√	—	√	—	—	√
1953	√	—	—	√	√	—
1954?						

We also find that there is a slight tendency to favour the non-preponderant outcome in the independent type of event (coin-tossing) to a greater degree than in the dependent event (e.g. sending presents).

By using a different technique suitable for youngsters we have attempted to discover how younger children of six and seven years of age predict in comparable situations. The situation may be briefly described as follows:

A display board with 16 pairs of lights (see Fig. 1) is shown to each child individually. This board is about 4 feet high and 1 foot wide. The pairs of lights are numbered 1 to 16 from the

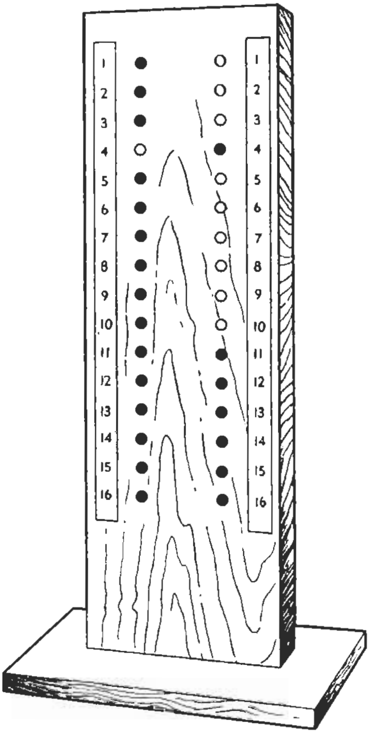

FIG. 1. Display Board

top of the board to the bottom. One light in each pair is lit in the first ten pairs, nine lights being on one side of the board and one light on the other. The lights appear rapidly one after the other and remain lit.

The children are told in advance that only one of each pair of lights will go on. After seeing one light in each of the first ten pairs lit, each child has to guess on which side the *eleventh* light will appear. He makes his choice by pointing to the small lamp in the eleventh pair which he expects to be lit next. The single light which appeared on one side of the board, whilst the nine appeared on the other side, was not always in the same pair of lights. It varied in random order from the first pair to the tenth pair in different experiments. The nine lights on the side of the board varied accordingly. For instance, if the single light was in the sixth pair and appeared on the left side of the board, the nine lights would appear in pairs 1, 2, 3, 4, 5, and in 7, 8, 9, 10 on the right side of the board.

We shall now consider the choice of right or left side for the eleventh light when the position of the single light is varied in successive experimental trials. The choices of children of six or seven years of age in this kind of binary situation seem largely determined by their choices in the previous trial. If they have previously chosen the right side for the light to go on, they will on the next occasion tend to choose the left side and so on. They do not seem to be influenced by the display of lights in front of them. They alternate their predictions from one outcome to another in successive guesses, merely oscillating their choice of side from right to left and left to right in successive guesses. The tendency to oscillate in this manner is weaker in the seven-year-olds than in the six-year-olds. To the extent that they do not merely alternate their choices from side to side, both age groups tend to predict that the eleventh light will be lit on the side with the single light, not on the side with the nine lights. In other words, they prefer to predict the non-preponderant side or outcome.

Between the ages of six and eleven years characteristic differences emerge in the way children predict the outcome of a binary event when it occurs in a series under conditions when they see whether their guess is correct or not. These differences can be studied with the help of the display board already described by means of a different type of experiment. Before any one of a pair of lights is switched on, the child has to guess which one of the two it will be. One of the first pair is then lit and remains lit. The child is then asked to guess which one of the *second* pair will next be lit. Again, one of this second pair is lit and remains lit. This is continued until one of each of the 16 pairs is lit. The particular lights which are lit on each side follow a pre-arranged scheme such that eight lights eventually appear on each side in random order.

Characteristics of Prediction at Different Stages of Development

Let us now follow the changes in the kind of prediction. From six to seven years of age children tend to alternate their guess after succeeding in the preceding guess, but not after failing in the preceding guess. After an unsuccessful guess they seem to show no special preference for predicting an outcome on either the left or on the right side. They are also inclined, though to a less extent, to switch from the side of the previous *outcome* and from the side with more *outcomes*.

Children from about seven to eight years of age differ from the younger ones in that they do not merely switch their guesses to the light on the other side of the display board. Alternation does not occur for its own sake, as it were. They are also unlike the six-year-olds in not being influenced by success in their previous choice. They tend, however, more strongly than the six-year-olds to change from the side of the previous outcome and the side with the preponderant number of outcomes. On the whole they are much more under the influence of display of the outcomes in front of them than are the children a year younger.

Ten-year-olds, on the other hand, do not switch their guesses merely because they have succeeded in their preceding attempt to guess. Their guesses are much more affected than those of the younger children by the number of lights that have already appeared in front of them on the display board, and the effect is such to make them favour the non-preponderant light.

Two features are characteristic of the entire age range six to eleven years. First, the tendency to alternate from the side of the previous light: if the previous light is on the right, the next choice tends to be on the left, and *vice versa*. Secondly, the children are not affected by what they have previously *guessed* but by the previous *outcome*, that is, by the *side* on which the previous light appeared, as well as by success in the previous guess and by the number of lights which have appeared already, as stated above. In the preceding experiment in which the children were shown ten lights and had to guess the side on which the eleventh would appear they (at least the younger ones) were affected by what they had previously guessed.

In the experiment just described there were eventually eight lights on each side in random order. We may now ask what would happen if there were consistently more lights shown on one side than on the other. It might be expected that, as the lights tended to be lit more on one side than on the other, there would be an increasing tendency to guess that the next light would appear on that side. In the extreme case, when all the lights appeared on one side we should expect the children to favour that side consistently, after a certain number of guesses. What we actually find is this: if the lights are lit so that eventually there appear 12 lights on one side and three lights on the other, the guesses for early lights are not different from the guesses when the lights are evenly shared by both sides; here, as there, there is a tendency to predict that the next light will appear on the side with fewer lights—the non-preponderant side. In the later lights, however. a preference gradually appears to predict the side on which *more* lights have already appeared.

Luck, Magic and Causality

We now turn to describe experiments the results of which will enable us to see more clearly what influences our predictions of outcomes which are independent. The subjects in these particular studies were large numbers of girls and boys, aged 12 to 15, in Grammar Schools.

The procedure is briefly as follows: a series of similar situations is described in which there are two possible outcomes and one of these has repeatedly occurred on successive occasions in the past. For example:

> Joan and Mary, two captains of school teams, are about to toss a coin before a match. Joan has won and Mary has lost the toss in the last eight matches. Who do you think is more likely to win this time? [1]

The number eight (matches) appears in some forms of the question. In other forms, any number from one to ten is substituted.

[1] The order of the names Joan and Mary is interchanged in some forms of the question.

An alternative form of the question is as follows:

> A girl is tossing a penny. The coin has turned up head on all of the last eight tosses. What do you think will turn up the next time it is tossed?[1]

We found that an appreciable proportion of subjects aged 12 to 15 years (roughly 15 per cent. of the younger ones and 27 per cent. of the older) do not say that *Joan* will win or that *Mary* will win. And if they are answering the question in terms of 'head' or 'tail', they do not say that '*head*' will turn up or that '*tail*' will turn up. They say that *either* Joan *or* Mary will win. Or they say that *either* head *or* tail will appear. Of the remaining 85 per cent. (or 63 per cent. of the older ones) of the subjects more than half say that the next winner or outcome will be *different* from the previous run. The reasons which the subjects give for their choices permit us to understand why they do not invariably predict that outcome which was previously non-preponderant. Nearly half the subjects stated that they believed it was time for the previously non-successful winner to win or previous outcome to occur, or words to that effect. But often they adduced other reasons. They said, for instance, that the tosser was skilful and so the run of her successes would continue. A rich variety of reasons is given which are often of a highly composite character. The prediction or guess is quite clearly far from being a simple one in most cases and varies considerably from person to person.

The reasons or explanations given by the subjects have been classified so that those based on ideas of dependence are distinguished from those based on ideas of independence.

[1] In some forms of the question 'tail' is substituted for 'head'.

The frequency with which each type of reason is given is tabulated below.

Table 2. Frequency of reasons given for predicting the outcome of a binary event

	Age	
	12–13[1]	14–15[2]
Dependence		
Favouring Change	24	24
Luck	23	11
Determinate	16	18
Series	6	8
Magic	3	4
Total	72	65
Independence	13	24
TOTAL	85	89

[1] Number in sample = 64 [2] Number in sample = 57

Some subjects gave more than one reason, so the total number of reasons shown in Table 2 exceeds the size of the sample. Typical explanations are quoted in the Appendix to this chapter.

One of the main features that emerges from the above classification of reasons is the relatively high proportion of the younger age-group, as compared with the older, who give 'luck' as the reason for their prediction. A second feature is the relatively small proportion of the younger group who say that the outcome cannot be known.

Perhaps the simplest way to interpret these results is to consider them as treating the unknown outcome as *dependent* on or as *independent* of previous outcomes. It will be recalled that the proportions of these two types in the total explanations are about 85 and 15 per cent. in the younger group and 73 and 27 per cent. in the older group of subjects.

Among those who justify their prediction in terms of *dependence* we find two antithetical elements of explanation. Some subjects speak of dependence in such a fashion that it leads to *continuance* of the *same* outcome. Others treat dependence so as to lead to a *change* of outcome. The former group regard the information about previous outcomes as telling them about the personal characteristics of the tosser, her skill, her determination or her luck. Alternatively, previous outcomes tell about the characteristics of the coin. This information they extrapolate and apply to the next throw. The latter group also believe the unknown outcome to be determined by previous outcomes, but not because these previous outcomes reveal the characteristics of the tosser or the nature of the coin. The mere fact that a particular outcome has occurred one or more times or that one person has had a run of luck means for them that there will be a *change* of outcome. In the simplest form this is expressed in the idea of alternation, for example, "I believe Mary will win next time because it goes alternately."

Among those who justify a prediction in terms of independence, some place emphasis on the indeterminacy of the event, foreknowledge being therefore impossible. A few seem to suggest that though the outcome itself is determined, no one can ever know in advance what it will be: "It's just up to the penny."

Stages in the Idea of Independence

We can now attempt to trace the development of the idea of independence in so far as it emerges in these experiments on binary events. At the youngest age which we have investigated, six to seven years, the most outstanding characteristic is a tendency to alternate from the previous choice. To the extent that they do not merely alternate their guesses they are apt to choose that outcome which was in the past non-preponderant. From their remarks in reply to questioning it seems that they think this is fair. Whether they have succeeded or failed in the last attempt also influences their next guess; after success they alternate, after failure they do not. To a comparatively slight extent they tend, in addition, to alternate from the side of the previous *outcome* and from the side with more outcomes. In short, the six-year-olds seem to be powerfully inclined to alternate their predictions as a result of a variety of specific influences.

Children a year older are much less apt merely to alternate from their previous choice. Like some of the six-year-olds but to a more marked degree, the seven-year-olds tend to change from the side of the previous outcome and from the side with more outcomes to the side with less outcomes. But unlike the younger ones mere success or failure in the preceding choice does not influence their next guess. They are much more under the influence of the display of outcomes facing them.

Most children, at least until the age of 10+, do not at first distinguish between an event, such as coin-tossing, the outcome of which is in reality independent of previous outcomes, and an event, the outcome of which is in varying degree dependent on previous outcomes, such as the state of the weather (fine or not fine) in a given place on a certain day. Predictions made about these two types of situation are similar for short runs of, say, eight events. The non-preponderant outcome is favoured even when the results actually show the children that one outcome has occurred three times as often as the other. Children may treat such short series of independent and dependent events without distinction as a single unit, and so favour the non-preponderant outcome. This effect is very marked in series consisting of five or six events.

In longer series, however, they do seem to make the distinction. Under such conditions when the event is of the coin-tossing kind, they predict in much the same way as they do in shorter series. When, however, as with the display board, the outcome in the earlier part of the series may be thought of equally as independent or dependent, the predictions in the later part of the series (after about eight outcomes), favour the outcome previously preponderant. This happens when the ratio of preponderant to non-preponderant outcomes is 3 : 1, and would no doubt be even more marked with higher ratios.

An idea of independence expressed in the phrase "either may happen" does not appear in our experiments at the age of ten. It first seems to emerge at about the age of 12. Thereafter the tendency to favour the non-preponderant outcome weakens and the inclination to give judgments of "either" is strengthened.

Decisive changes in the idea of independence take place between the ages of 12 and 15, at least in the minds of some adolescents. This idea, which begins at the age of 12, is clearly and forcefully held by some boys and girls at the age of 14.

The notion of luck has a variety of forms and meanings at this period. Gradually it merges with the idea of chance in the sense of independence. For instance, a girl says: "It is just luck." Even at this age, however, the vast majority entertain ideas of dependence. These appear in two opposing forms, which are sometimes held by the same boy or girl: the one form of dependence favours continuance of a series of similar outcomes; the other form favours a change to the other possible outcome, as the series proceeds. The belief in luck is still potent and faith in magic is by no means absent.

From a variety of experiments of the kind described above we are able to trace some of the rules of subjective probability throughout the years of development. These rules vary characteristically from year to year and differ strikingly from the rules of mathematical probability. It would hardly be an exaggeration to say that there is no such thing as a 'guess' in a sheer random sense. Apparently, from the youngest child onwards, predictions and judgments follow an intricate system of preference and pattern-seeking varying in complexity from age to age. On the basis of this system the child makes his choices. He interprets what has happened, decides what to expect, and makes up his mind what to do.

The grounds for these preferences may be in past encounters with different kinds of events. Most children will doubtless have had experience of binary events the outcomes of which seem to them to occur more or less alternately. We might say that the youngest children tend to attempt to express their understanding of independence by giving judgments of alternation. When a child of six or seven predicts alternate outcomes this is his way of saying that the outcomes are independent. Children also meet, and to a much larger extent, the kind of event which has a recurring outcome. Indeed, much learning takes place in relation to such recurring outcomes. In some of our experiments, the children soon learn that the events are of the alternating type. In other experiments, they have to use their experience of similar events in order to decide whether the experimental event is of the alternating or recurring type. When, for example, a child is told that a coin has appeared head eight times in successive tosses, he might regard the outcome as a recurring one if he had not had experience of coins behaving in a non-recurring fashion.

As childhood passes into adolescence, the systems of choice based on privately formed influences weaken and tend gradually to become displaced by systems which are objectivley determined or which are inherent in the objective situation. A child of six, when faced with the task of predicting the outcome of a binary independent event, will tend simply to alternate from his previous choice. A child of nine or ten will tend to choose the outcome which has happened less often previously. Some children of 12 can already begin to recognize in events of this kind that the outcome on any occasion is entirely unaffected by what has happened on other occasions. Few adults have completely outgrown the features which characterize judgments at the various earlier, stages of intellectual growth.

Note

The tendency to regard separate independent events as part of a single continuous series, although we "know" they are independent, we propose to call the phenomenon of Subjective Pseudo-Dependence. This tendency, previously observed at the roulette table and noted by Poincaré, cannot be dismissed as an oddity or shortcoming of the mind manifested only in gaming. It is a remarkable and deep-seated characteristic which has been clearly demonstrated in our experiments. In this chapter we have begun to trace its development and to identify some of the factors which influence it. We shall encounter it again in the study of sampling and of gambling.

Appendix

Some typical reasons for the choice of outcomes are given below:

Favouring change

"Because if Mary has won seven times she is not likely to win again."

"Mary (will win) because Joan won all the other times and it is about time Mary won."

"I think Mary will win because on the Law of Averages it is Mary's turn to win."

"Joan will probably win the toss because usually the one who lost the time before will win the next time."

"Because usually the one who has been losing a lot wins. But really it is just luck."

Luck

"One girl has won four times but her luck has run out so the second girl wins and the first girl loses because tails has not won." (This also favours change.)

"Mary seems to be the lucky one of the two and she won first time."

"Joan seems to have all the luck so I'm sure she will win again."

"I think it would be tails because most people think it is the luckier of the two."

Determinate

"Because Mary most likely can only throw a certain way and Joan knows what it usually comes to."

"Mary won the last two so she must be the best and if she is the best she will surely win the next one."

"Because if Joan has won six already she is obviously better than Mary and so it is possible that she will know more the next time."

"Because the wind might be causing the coin to drop on the side that it has been doing previously."

"Joan (will win). Joan would be very determined to win again and if the same person tossed she would probably call what she called the last time which was heads. Mary would be very nervous and have given up hope of winning."

Series

"Tails. Because it has turned up tails three times already and it is sure to turn up the same after three times turning up tails."

"I think heads will turn up next time because it has for five times."

"It will turn up heads because if it has turned up the same ten times before it will probably do so again."

"Mary is likely to win because the coin usually turns up six or more times the same, if it has turned up five."

Magic

"Mary will think she will win and so probably will not win."

"Heads. Because for some strange reason, the fifth toss always comes heads. I read this in a book, and tried it often, and it always works (book title, 1000 wonderful things)."

"If the girl chose tails I think it will be tails because there is a saying that 'Tails, tails never fails.' This saying is often very true."

"Tails. Because seven is my lucky number and it is the seventh time it has been tossed after six heads. So I think it will change to tails by the seventh time." (This also involves the idea of luck.)

Independence (Indeterminate)

"It depends on the way the coin is tossed up. Any one of them might win. Mary has won the guesses but this does not mean that she will win this time. She might very well but Joan could win this time. It is not possible to tell who will win."

"Heads or tails. It's just up to the penny."

"Because if a penny spins in the air it does not necessarily land on the same side every time."

"Because a penny has two sides a head and a tail, and no one can tell which end it is going to land (unless they cheat) so it could be either heads or tails."

"Because when you toss a penny nobody, *nobody* can tell even the cleverest of people which it is going to be, because it might be either."

"Either. It just depends on which side it lands after being tossed."

Chapter Three

SAMPLING THE UNIVERSE

The reader might like to begin this chapter by undertaking a little experiment in sampling and then testing the accuracy of his estimate with the correct figures. Let him turn to page 20, count the number of times the letter e occurs in the first line, and then without looking at the rest of the page, let him write down how many e's he thinks there are in a page as a whole. Now let him count the number of e's in the first five lines and make a second estimate. After that, let him count all the e's on that page and compare his two estimates with the correct number.

We often have to make up our minds about a group of people when we have only met a few of them or about a person after a casual encounter. An employer interviewing an applicant, a professional tea or wine taster, a courting couple, a Congressman judging all Chinamen after seeing two, a publisher reading a specimen chapter of a book, a pilot taking up an aeroplane on a test flight, a reviewer skimming a table of contents, and a lady in a perfume shop, all demonstrate the use of sampling in everyday life. The list could be continued indefinitely. Sampling is not merely a convenient practice. It is a biological characteristic of the activities of living things. Our senses function by a natural sampling of the outer world. Indeed they cannot work in any other way. The manner in which we selectively attend to what goes on around us is only a more complex sampling device. We can only absorb a minute proportion of what impinges upon the senses. The nature of human experience is, of necessity, a constant sampling of the universe.

The things, people or situations about which we seek knowledge in any given instance may be very unlike or much alike. To the extent that they are unlike, the larger the sample the fuller the information we shall have about the population to which they belong. If they are substantially similar in respect of the quality in which we are interested, then a sample of *one* may tell us enough. If we are buying biscuits, the grocer may offer us a kind which we have never tasted before and we see that all the biscuits in the tin bear a close resemblance to one another. He gives us one to taste and we take it for granted that the others will be indistinguishable and that if we like or dislike this one we shall like or dislike the rest. So on the basis of a sample of one biscuit we form definite conclusions about the total contents of the tin and about this kind of biscuit in general. A mistaken conclusion about biscuits in general might not be very serious, but a young man, rebuffed on his first encounter with a young lady, might be discouraged from enlarging his sample of encounters with the other sex and might, by faulty generalization from too limited a sample, find himself a celibate for the rest of his life.

Of course in at least one very large sphere of activity we are in no position to choose the size or variety of sample we should like to have. We are forced to rely on a sample of very small size and limited in range, although the entire future of the countless individuals concerned may be at stake. We refer to the world of examinations and tests of all kinds in education, industry, the professions, sport and social life generally. An examination is supposed to be a sample of a candidate's ability, skill or knowledge. The difficulty which academic examiners have in assessing the merits of their students is notorious. Their private misgivings are not reflected in official results. Not every examiner seems to have such qualms. Some claim that they need not read the whole of a student's script; one page or even a paragraph may be enough for them to place the student in his proper class, and others allege that as good an order of merit of candidates may be obtained by weighing the students' scripts as by reading them.

The most expeditious method of examination is without doubt that adopted by a blind old Chinese priest when judging an essay submitted by a candidate for the Civil Service. Unwilling to waste his time listening to the essay read out to him, the priest politely asked if a few pages could be burnt so that he could smell them and so arrive at a considered view. He sniffed and announced that the essay had some quality and that the candidate would probably pass. The young man, wishing to test the validity of this somewhat unusual procedure, set fire to a classic, whereupon the priest in transports of delight proclaimed: "Ah, genius!" Whereupon the candidate, producing the charred remains of his own effort, saw with chagrin the old priest manifest growing repugnance until in the end he vomited.[1]

A sample can only tell us about the attributes of a population when these attributes are represented in the sample. But in everyday life it is not unusual for people to generalize about certain attributes in a population on the basis of quite different

[1] Based on a review of D. Woodruff's *Walrus Talk* (London: Hollis and Carter, 1954), which appeared in the *News Chronicle*, 29 December 1954.

characteristics in the sample. The Greek architect who hawked a brick around as a sample of the house he was going to build illustrates this practice. The extraordinary behaviour of the Englishman in France, divested as he is of local restraint, forms the basis for rather curious generalizations by Frenchmen about the British national character. Part of the Englishman's strange conduct may itself be due to his remarkable generalizations about Frenchmen (and Frenchwomen), based on clandestine delving into French novels and surreptitious visits to the more uninhibited music halls.

It is worthy of note that the process of 'sampling the universe' corresponds to what Freud held to be the essential nature of thought. He supposed that in the early stages of the development of the human mind, objects of desire were represented by hallucinations. This is the form they still assume in our dreams, which almost certainly employ archaic mental processes. We may put this in another way by saying that when the object of desire was not immediately attainable, it was represented consciously in the form of some more or less vivid image. The image, as Sir Russell Brain describes it, is "an enduring model of the external world".[1]

But, argued Freud, hallucinations, however life-like they are, do not bring the expected pleasure. We may conjure up in our imagination the most wonderful banquet but we shall remain hungry. Our appetites may even be whetted rather than appeased by the image of a sumptuous feast. So there evolved a new type of mental process based on a recognition of reality and not just, as earlier, on the memory of a pleasurable aim. Accordingly, it became possible for impulses not

[1] Brain, W. Russell, *Mind, Perception and Science*, pp. 44–6. Oxford: Blackwell, 1951.

to be released until conditions favoured their gratification. This enhanced the importance of the organs of sense. There came into being a capacity to sustain the attention, to search or scan the external world from time to time so as to become more familiar with its true nature.

The Freudian theory amounts to this: our conscious selves interpose between wishing and acting "the procrastinating factor of thought". In its very nature, thought is an experimental method which expends small quantities of energy. When we think, we are behaving in much the same sort of way as a general when he manipulates his little symbols on the operational map before going into action. Thought enables us to anticipate the likely consequences of an action by trying it out internally on a minute scale.[1] The apparatus of thought resembles the sensory process out of which it grew. Just as our perceptual system samples the outer world, so in thinking we send out sensitive feelers, groping forward whilst expending only a relatively small effort and retreating to assess the sample. Thinking is thus, economically, a great improvement over the preceding stage of impulsive action in which large amounts of energy were wastefully employed.

We have been using the word population in its usual sense to refer to an aggregate of things or people that physically exist at the time the sample is drawn. The word can also be used in the sense of a population that is generated, for example, by some mechanical device, because of the properties built into it, or by some human or animal subject. Imagine a situation in which a person has to press a switch which may or may not provoke a noise. Both the ratio of occurrence to non-occur-

[1] Freud, S., *New Introductory Lectures on Psychoanalysis*, p. 124. New York: Norton, 1933.

rence of noise and their order can be set in advance by suitable adjustments and the subject's predictions can be subsequently assessed.

Experiments in Sampling

The experiments now to be described are only first stages in what is a large and varied field of investigation on which we have recently embarked. The scope of these studies may be briefly sketched. Our initial task is to discover how people make estimates or inferences on the basis of a sample when the population from which the sample is drawn remains unchanged and the size of the sample is increased by equal steps. This might be exemplified in everyday life by an egg-merchant who is testing the quality of eggs by batches taken from a large consignment. Or by our Englishman returning from abroad with opinions about the succulence of French cooking, the lack of potency of French beer, and the exuberance of French nocturnal entertainment which grew firmer as he prolonged his Parisian sojourn.

Having explored how estimates vary with the size of the sample, we can go on, at the next stage, to study the effects of varying the composition of the population. We can introduce any ratio of the elements of the population and keep it fixed throughout the experiment.

A more complex form of sampling arises when the population is not fixed for the duration of the experiment but changes in a manner known to the experimenter and so affects the sample in corresponding fashion. For instance, we can add or remove blue beads from an urn in a systematic manner between each estimate about the contents of the urn made by a subject. The estimates based on the samples drawn from the urn will then reflect fluctuations in the population of the urn. We can study the subject's estimates of the population and, what is of particular importance, the way he interprets the fluctuations in the sample, arriving perhaps at some rough and ready natural law. Such enquiries might lead to a better understanding of the way regularities in natural phenomena are noticed in everyday life. The merchant who frequently sampled his consignments of eggs might form an idea that they were better in winter than in summer.

A naturally generated population might be illustrated by the height of the tide which varies in a systematic manner. Superimposed on its variation, however, are daily fluctuations due to an enormous number of influences and therefore, for the observer, apparently random. Nevertheless, a patient and persistent watcher of the tides might be able to see beyond diurnal variations and get an idea of the rhythm and periodicity of the basic movement.

It is also possible to study variations in the time element in sampling. The entire sample may be drawn in one batch or in successive batches or in single items. The time interval between batches or items may be varied by the experimenter.

A sample may be used to give two kinds of information. It may be used to predict what will happen next. Or it may be used to estimate the nature of the population from which the sample is drawn. In the previous chapter, on ideas of independence, we have shown how a sample is used by children of different ages to predict the outcome of an event. In the present chapter we are concerned with this sort of prediction when it is based on larger samples and we shall examine the relationship between the prediction of the next bead and the estimate of the population in the urn.

The experimental situation which we here study is one in which there are two possible outcomes of each event, but we can easily envisage a situation in which there are three or any other number of possibilities. It goes without saying that beads have been chosen for convenience in studying large groups. Other experimental content and methods could, of course, also be used.

We shall begin with the simplest type of situation. The sample, starting from zero, that is, before any bead is drawn from the urn, increases steadily one by one. The subjects are shown an urn and are told that it contains a hundred beads and that each bead may be either blue or yellow but no other colour. No information at all is given about the porportions of blue and yellow beads in the urn. One bead at a time is taken from the urn and held in the hand without being shown to the subjects, who are divided into two groups. One group guesses the colour of the bead which has been drawn. The other group guesses the number of blue and yellow beads in the urn. As the sample is enlarged one by one, both groups of subjects make a fresh prediction or estimate. Each group makes 20 estimates, the first when the sample is zero, and the last when the sample is 19, that is, when the subject has seen 19 beads drawn from the urn and replaced.

We have carried out this experiment with boys aged 12 and with an adult group consisting of staff and students at Teachers' Training Colleges.

The predictions and estimates made by the subjects may be set out, as in the two tables which follow, and compared with the actual percentage of blue beads drawn at each stage.

From a comparison of the two sets of estimates, made by the twelve-year-olds and adults respectively, we can attempt to answer a number of questions. The reader's attention should be drawn to the fact that in drawing beads singly from an urn,

Table 3. Sampling: Predictions and estimates made by twelve-year-old boys

Colour of bead	Percentage of blue beads so far drawn (Sample)	Percentage predicting blue bead	Mean percentage of blue beads in urn as estimated
—	—	75	67
Y	0	50	50
B	50	42	62
Y	33	83	54
B	50	25	56
Y	40	83	59
Y	33	83	47
Y	28	67	49
B	37	35	56
Y	33	42	44
Y	30	67	40
Y	27	67	47
Y	25	42	38
Y	23	58	40
B	29	42	53
B	33	50	57
B	37	50	49
Y	35	67	50
Y	33	58	43
Y	31	75	42

whatever proportions of blue and yellow beads we put in, we cannot guarantee that these proportions will be even approximately represented in a sample of 20 beads. In fact, although the ratio of yellow to blue beads in the urn, in both experiments, was three to one, in the experiment with adults only two of the 19 beads drawn were blue and in the experiment with twelve-year-olds, six were blue.

Table 4. Sampling: Predictions and estimates made by adults

Colour of bead	Percentage of blue beads so far drawn (Sample)	Percentage predicting blue bead	Mean percentage of blue beads in urn as estimated
—	—	65	59
Y	0	47	44
Y	0	72	52
Y	0	72	42
Y	0	35	31
Y	0	44	21
Y	0	33	16
Y	0	19	13
Y	0	16	9
Y	0	14	5
Y	0	14	5
B	9	33	8
Y	8	19	9
Y	7	12	9
Y	7	21	8
Y	6	21	6
B	13	26	9
Y	12	23	13
Y	11	21	13
Y	11	30	10

(i) What is the effect, if any, of increase in the size of the sample (*a*) on predicting the colour of the next bead and (*b*) on estimating the proportions of beads in the urn?

(*a*) As the sample increases there is a tendency on the part of the adults to predict at first the bead colour which has appeared less often before and then the colour which has appeared more often. The twelve-year-olds fluctuate more in their predictions and many of them favour the non-preponderant colour in the next choice (as in the experiments on ideas of independence). We should note that the adults had an easier task in that the first ten beads drawn in the experiment with them were all yellow. It should be particularly emphasized at this point that variations in the composition of the sample constitute perhaps the most important variable in experiments on sampling. Compare two extreme cases: one in which the sample is completely homogeneous, e.g. all blue beads; and the other when the proportions of blue and yellow are equal but the order of appearance of blue and yellow beads is random. In the former instance additions to the sample do not alter the proportions. In the latter instance, additions make a maximum change.

(*b*) When the first bead is drawn, whether it is blue or yellow, it cannot contradict the assumption that there are equal numbers of blue and yellow beads in the urn. We find that two-thirds of the adult subjects assumed a 50 : 50 ratio of blue and yellow before the first bead was shown, and that after they saw what colour it was, not one of them altered his estimate. The remaining subjects did, however, change their estimates. The average estimate of the adults tends to approach closely the proportions in the sample. The estimates of the twelve-year-olds show only a slight tendency to approach the proportions in the sample. Their first estimates favoured blue, but the later ones favoured yellow, in accordance with what appeared in the sample.

(ii) What correspondence, if any, is there, as the sample increases, between the estimates of the contents of the urn and the predictions of the colour of the next bead?

The estimates and predictions given by the adults are much closer than those given by the twelve-year-olds. As Figs. 2 and 3 show, the proportion of blue beads estimated to be in the urn by the adults is very similar to the proportion of

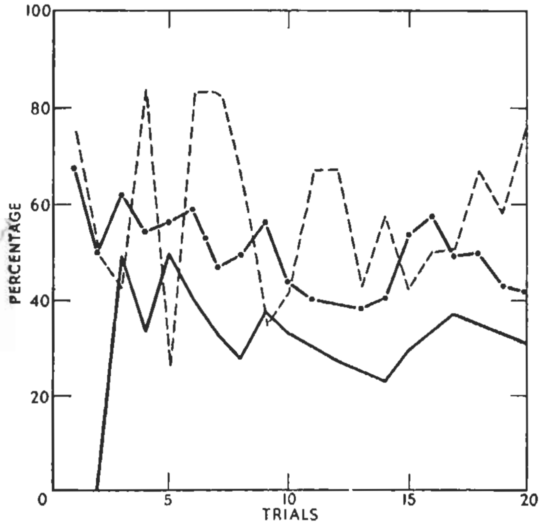

FIG. 2. Sample, predictions and estimates (adults)
.—.—. mean percentage of blue beads in urn, as estimated ($N = 9$).
- - - - - percentage predicting next bead blue ($N = 43$).
——— percentage of blue beads in sample.

adults predicting that the next bead will be blue at each successive bead drawn. The twelve-year-olds give estimates which differ markedly from their predictions: the proportion of blue beads which they estimate to be in the urn is not at all similar in general to the percentage predicting that the next bead will be blue.

Here we meet an apparent paradox, the inconsistency between the predictions and estimates of the twelve-year-olds. In their predictions they favour the non-preponderant colour

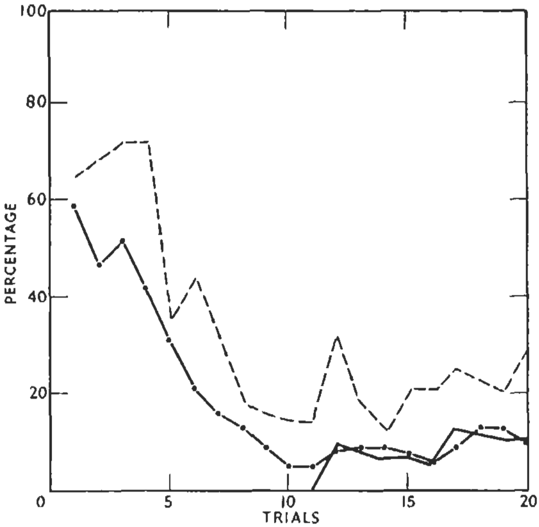

FIG. 3. Sample predictions and estimates (age 12 +)
.—.—. mean percentage of blue beads in urn, as estimated ($N = 9$).
- - - - percentage predicting next bead. blue ($N = 9$).
——— percentage of blue beads in sample.

and in their estimates of proportions in the urn they are closer to the preponderant. We can explain this curious feature if we recall an influence which we have encountered previously—the tendency to favour the non-preponderant outcome. When there is a run of three yellows there is an increasing tendency to predict blue. An inclination to favour a non-preponderant outcome cannot arise in estimating proportions in the urn. Most of the adults, however, do not compensate or balance in this way and predict more or less on the basis of the objective information given, rather than in terms of subjective preferences.

(iii) Suppose in this experiment we had placed beads of only one colour (e.g. blue *or* yellow) in the urn, how would the estimates be affected by the regular appearance of a bead of one colour?

Although we have not carried out an experiment under these conditions, it so happens that the first ten beads drawn in the experiment with adults were yellow. As stated above, two-thirds of the adults began by estimating equal numbers of blue and yellow beads in the urn. This proportion of adult subjects steadily diminished until after the sixth bead had been drawn, no one estimated a 50 : 50 ratio. In the first four sets of estimates not a single adult estimated that all the beads were yellow. From the fifth to the ninth estimate, the number of adults estimating no blue beads in the urn gradually increased until after seeing eight beads, five of the nine subjects estimated all yellow and no blue beads. Even after nine yellow beads had been successively drawn, four of the adults still believed that there were some blue beads in the urn.

(iv) Does confidence in predictions or estimates vary as the sample increases? Is a person more confident of his prediction or estimate after seeing a sample of ten than after seeing a sample of five or one, regardless of the composition of the sample? If two samples have the same composition, does the larger sample give greater confidence?

In the present experiment we have no direct measure of confidence but we can indirectly measure it by the amount of change in each subject's estimate after he has seen an additional bead drawn from the urn.

We find that estimates successively change by relatively large amounts at the start, when the sample is small, and change less and less as the sample increases in size. This effect is doubtless bound up with the composition of the sample. In our experiment with adults, after the tenth trial, when the first

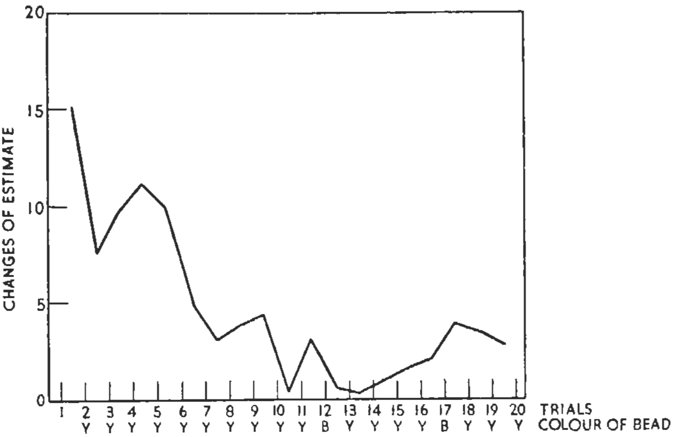

FIG. 4. Changes in successive estimates in relation to size of sample (adults $N = 9$).

blue bead was drawn, the successive estimates began to fluctuate rather more widely, though by no means as much as at the outset. (See Fig. 4.)

(v) Is the estimate based on the entire sample of beads drawn or is it more influenced by the beads which have last appeared?

This sort of problem arises often in everyday experience. We tend, on the whole, to be more influenced by recent events in a series. This may be due partly to the fact that the later members of the series are fresh in the memory and the earlier ones forgotten, and partly because we are sampling a changing population, which may be better represented by later outcomes. It would normally be safer to predict a cricketer's next score on the basis of his ten most recent innings rather than on all the past scores he has ever made. In the same way, it is often wiser to attach more importance to a student's efforts in his final year than to his vagaries in the first term.

In the present experiments it appears that the younger subjects are much more influenced by recent outcomes than the older subjects. The estimates of the contents of the urn made by the twelve-year-olds most resemble the composition of the sample when we limit the sample to the most recent three to six beads drawn.

All these questions and answers have been expressed in terms of group means but we cannot be certain that such averages fully reflect individual types. If, to take an extreme instance, half the subjects estimated 100 per cent. blue and half estimated 100 per cent. yellow, an average would give a 50 : 50 ratio of colours. In Chapter Two, we have attempted to demonstrate type differences which would be blurred by averages alone. In the present experiments we have not yet enough information about individual subjects to enable us to deal with this question adequately. It would seem, however, that, bearing in mind the differences in the samples drawn, the adult estimates seem to vary much less than those of the children among themselves.

We can now consider very briefly the place of sampling in everyday experience, particularly in appreciating regularities in natural phenomena and in forming impressions of other people. Students of learning have been mostly concerned with fixed situations which have an unchanging series of events, and with changes in reward or punishment rather than with changing events. In daily life, however, we meet not regular sequences of events which are all the same, but irregular sequences, and the irregularity cannot usually be foretold. We may have to make frequent contact with a person whose moods are unpredictable. They cannot be foreseen and we have to learn what to expect of that person as he appears in the office in the morning. A cricketer has a flaw in his batting technique; when he makes this error he is not always out. Whether he is out or not depends upon the bowler, fielders and other factors in the situation. He is more often out when he makes the mistake than when he does not make it, and his task is to realize in fact, that he is more often out when he commits the error than when he does not, and to learn to eliminate it under the diverse conditions under which it occurs. Exactly the same circumstances occur in many other situations, not only in sport, where there are random influences which are not appreciated by those concerned.

Learning in childhood is very largely a form of sampling. A child constantly encounters uncertainties of one kind or another which may lead to instability. He has to learn to interpret these uncertainties so as to arrive at a realistic expectation of what is likely to occur. It seems important to eliminate as far as possible random and unpredictable elements in the child's relationships with parent and teacher. The child should know more or less what to expect and when to expect it, and in so far as situations occur in a child's life with outcomes which are not consistent, he must learn to recognize and cope with them. There are situations too, in which it is desirable to avoid rigidity and obsessiveness in the child's expectations. He should be allowed a certain freedom of manœuvre without having to anticipate a quite definite outcome. He should not expect disaster if a usual outcome does not materialize. He should learn that situations are not normally irretrievable and that severe punishment is not an inevitable consequence of a change in outcome.

Chapter Four

THE LANGUAGE OF UNCERTAINTY

Mrs Jones remarks to Mrs Smith: "My husband is always coming home drunk." What precisely does Mrs Jones mean? Does she mean that her husband is invariably intoxicated when he crosses the threshold? If not, what proportion of sober homecomings has Mrs Jones chosen to overlook? Language is rich in words like 'always' which can have a numerical meaning within a fairly wide range of values. All such words or phrases which denote magnitude, intensity, frequency, duration or likelihood can have such meanings. Situations in everyday life call for considerable use of such expressions, and any newspaper is plentifully supplied with them. The following passage illustrates the use of words of this kind. These are italicised.

> A *long* preparation for a job, whether in the form of general education, professional training, or a period of *low* earning before a practice could be established, meant in the *past* the possession of *money*, and derived further prestige-value from that fact. But it *seems likely* that even if the public *eventually* comes to provide the money for everybody's occupational training there will still be a *tendency* to esteem the jobs needing the *long* training and the trainees on whom it *is thought* to be worth while spending *so much* money. The assumption is that the longer the training the more difficult the job *must be* and the further it is beyond the *ordinary run*

> of people. To have its full prestige-value the *long* training must precede paid occupation in the work, *no doubt* because this implies the spending of money on the learner. *Many* nurses in England have thought that the social standing of their profession would be *higher* if they had "student status" instead of being employed during the *long* period of their training, and this view, whether sound or not, illustrates the fact that length of training is *widely* held to earn respect for an occupation.[1]

In considering the words printed in italics in the above passage we raise the question whether it is possible to give them any quantitative meaning. Of course, when we use words of this kind we do so just because, in most instances, we cannot give a precise quantitative value. Indeed the apparent imprecision may give more information than would a single value. But we may still ask the user two questions: first, within what limits is the value of the word intended to lie? Second, what is the mid-value of the range to which it might apply? We can then see whether the listener or reader gives the word the same interpretation. If the author of the above passage had to give a single quantitative value for each of the italicised words, what would be the one he would choose? And we may wonder whether this value is the one which the reader has in mind, and also whether different readers have different values in mind, and finally, what is the range of values in the minds of different readers. For example, when the author says 'a long preparation for a job' does he have one year in mind, or ten years? Or when he talks about 'a period of low earning', does he think in terms of £2 a week, or perhaps £10 a week? Again when he uses the

[1] Harding, D. W., *Social Psychology and Individual Values*, p. 92. London: Hutchinson's University Library.

expression 'it seems likely', we may wonder what sort of probability is at the back of his mind. We might gauge his level of probability by the amount he would be prepared to stake on the issue in question. Or when he uses the phrase '*many* nurses', what absolute number does 'many' denote, and also what proportion of the total number of nurses does 'many' imply?

We have, in fact, obtained from a group of 27 adults their interpretations of some of the italicised words in the above passage. A tabulation of their replies gives us some idea of the range of meanings among different readers.

	Range of values
long preparation for a job	2 to 10 years
a period of *low* earning	£100 to £500
possession of *money*[1]	£250 to £75,000
eventually	3 months to 100 years
tendency	10 per cent. to 95 per cent.
no doubt	20 per cent. to 100 per cent.
Many[2] nurses in England	25 per cent. to 99 per cent.
widely held[3]	10 per cent. to 99 per cent.

The corresponding interpretations given by the author of the passage are:

long: longer than the average needed for the jobs with which any given job is being compared.

low: too low to support the entrant to the occupation in the material standard of living that his family and friends consider minimal for a man of his age.

[1] Some persons wrote £2000 per annum.
[2] Some persons gave absolute numbers, e.g. 1,000 or 15,000.
[3] Here also some persons gave absolute numbers, e.g. 35 million.

money: unearned income insufficient to bridge the gap between his earnings and the income needed for the standard of living indicated above.

eventually: this word is misused and should be replaced with 'ever'.

tendency: this word is misused and the sentence needs recasting: '. . . occupational training, the jobs needing a long training . . . will, other things being equal, be esteemed the more highly.'

no doubt: probably (with propagandist insinuation of a certainty that cannot be openly claimed; a misuse rapidly becoming an accepted idiom).

Many: enough to ensure that when the topic is raised with a random sample of about ten ward sisters the view can be alluded to and discussed with them without its having to be explained as a new idea.

widely held: met with (or implied) often enough to make us feel doubtful about securing a majority of our group against it if a vote were taken (e.g. for the last few years it has been a widely held opinion in Britain that the Conservative party gives us the best government and also a widely held opinion that the Socialist party does).

Experiments on Meanings

We have studied the quantitative meaning given by children of different ages and adults to three categories of word or phrase. The subject is shown sentences containing a word or phrase which denotes some quantity, frequency or duration. He then writes down what he thinks is a suitable numerical value for the phrase. For instance, after the sentence: 'There are a lot of trees in the park,' he answers the question: 'How many trees do you think there are in the park ?'

Five grades were used for each category and each grade was exemplified in three different contexts as follows;

Quantity	*Frequency*	*Duration*
Hardly any	Very rarely	In a moment
Not many	Seldom	Soon
Some	Sometimes	Not long
Many	Often	For some time
A lot	Nearly always	For a long time

The contexts for Quantity were :
Friends, Trees, Stones;

for Frequency, they were :
Raining, Late for School, Having a Cold

and for Duration they were :
Raining, A person being away, Time until Christmas.

We give below nine examples of the contexts as used in each category. In the actual experiment, the total number of sentences was 45.

Peter has many friends
There are a lot of trees in the park
There are some stones on the path

It is nearly always raining in Patagia[1]
Sheila is seldom late for school
Mary sometimes has a cold

It will stop raining in a moment
He will be away a long time
It will soon be Christmas

After each sentence there was a direct question to elicit the quantitative meaning. For instance, after "Peter has many

[1] Orally explained as an imaginary city.

friends" there was the question: "How many friends has Peter . . .?"

We used varying contexts because it seemed obvious that the words and phrases would have widely differing values when referring to different things. Compare, for example, the two sentences: "There are a lot of people in China" and "There are a lot of people in the room." Clearly the expression "a lot of" cannot be given any numerical meaning out of its context.

A first glance at our results suggests that the adult replies are more clear-cut on the whole than those of the children, so it would perhaps be better to look at the adult results first and then see to what extent the children's interpretations at different ages approach those of the adults. We shall first study the adult values given in the three contexts for the category 'Quantity'. The median values given by the adults are shown in Table 5.

Table 5. Category: Quantity. Numerical values[1] *given by adults*
$N = 27$

Grade	Context		
	Friends	Trees	Stones
Hardly any	4·3	9·3	6·6
Not many	2·6	13·0	13·5
Some	5·2	20·5	18·7
Many	15·5	138·0	175·5
A lot	15·2	150·5	175·5

[1] Medians.

It is clear that the phrases for the different grades vary in their numerical value with the different contexts. A person says there are hardly any trees in the park. When pressed, he says that he meant there are about nine trees in the park. If he says he has hardly any friends, the number he has in mind is about four. The replies given to the questions about stones and trees are remarkably similar, which suggests that the implicit numerical meaning of the phrases is fairly[1] stable in similar contexts. In two of the grades, 'not many' and 'some', the values for trees and stones are practically identical; in the three other grades the biggest difference in value between the two contexts is about 40 per cent. The influence of context is shown characteristically in the different values given in the context 'friends'.

We should perhaps note that even in one and the same context a particular phrase might be open to several interpretations. In reading the sentence 'there are some stones on the path' one person may think of a very stony path, another person might think of an asphalt path with a few stones dotted about. We should also note that we cannot expect equal differences between the values allotted to the different grades, for these were arbitrarily chosen. It is true that we expected that they would lie along a continuum with roughly equal differences between them, but in point of fact not all the subjects treated them as increasing in order of magnitude. 'Not many friends' is given on the average a smaller value than 'hardly any friends', whereas we had expected the reverse to appear; and 'many friends' and 'a lot of friends' are given

[1] If the reader cares to ask what degree of stability is intended, we venture to refer him to Table 5 where he will find a partial answer.

the same value, just as 'many stones' and 'a lot of stones' are given the same value.

It is possible that the *relative* difference between any two grades is affected by the context in which they appear. In the context of Frequency, the upper limit of possible values was fixed. The questions were always worded to ask how frequently the event occurred in a month, and in most cases the

Table 6. Category: Frequency. Numerical values (per month) given by adults
$N = 27$

Grade	Context		
	Raining	Late for school	Having a cold
Very rarely	1·1	—[1]	0·8
Seldom	1·8	0·9	1·0
Sometimes	5·6	3·9	2·8
Often	17·1	13·6	6·9
Nearly always	21·6	18·0	16·3

[1] This grade was omitted in this context.

event could not happen more than once a day. In this category the three contexts, 'having a cold', 'raining' and 'late for school', do not seem to have a very marked effect on the grades in the range of values given from smallest to largest. Nor do they affect the grades very much in respect of the *relative* differences between the grades. The ratio of 'seldom' to 'sometimes' is roughly about a quarter to a third in the three contexts, and a similar constancy is noticeable in the ratio of 'sometimes' to 'often' in the three contexts.

In the category of Duration, context seems to play an enormous part in determining the range of values allotted to the different grades and also to the relationship between grades. In the context of 'raining', the values in the grades range from about 15 seconds to about 2½ hours. In the context of 'a person being away' the value for 'in a moment' is about

Table 7. Category: Duration. Numerical values (hours) given by adults
$N = 27$

Grade	Context		
	Raining	A person being away	Time until Christmas
In a moment	0·005	0·004	0·002
Soon	0·033	0·040	28·500
Not long	—[1]	—[1]	29·850
For some time	1·460	137·000	81·100
For a long time	2·380	1307·000	271·500

[1] This grade was omitted in this context.

12 seconds and for 'a long time' 1307 hours (or eight weeks), and in the context of 'time until Christmas' the corresponding scale of values is from six seconds to 271 hours (or about 11 days).

Let us now consider the values given by the children. A discrimination of the various terms denoting Quantity seems to come at an earlier age than the discrimination of expressions for Frequency. The effect of an increase in age on the child's understanding of Quantity seems to depend on the context.

It shows itself in relation to 'stones' but not in relation to 'trees' or 'friends'. 'Some trees' means much the same number (13 or 14) at each age. 'Some friends' means about seven at each age but 'some stones' is given the value 14 at age seven, 19 at age eight, 54 at age nine, and 75 at age ten. Similarly, the adult's interpretation of Quantity depends on context. (See Table 8.)

Table 8. Category: Quantity. Ratio $\frac{\textit{a lot}}{\textit{hardly any}}$ based on values given by children and adults.

Context	Children, average age				Adults
	7·1	8·4	9·4	10·1	
Friends	4·5	3·9	4·4	3·7	3·5
Trees	3·0	3·7	4·5	4·3	16·2
Stones	2·8	5·1	5·5	6·2	26·6
Average	3·4	4·2	4·8	4·7	15·4

A slight effect of age is shown in the increasing values of the ratio $\frac{\text{a lot}}{\text{hardly any.}}$ All our subjects, even the youngest, give a value for 'a lot', which is, on the average, three to five times as great as the value they give for 'hardly any'. To the adult, however, on the average, 'a lot' means 15 times as much as 'hardly any'. In the context 'friends' the ratio is much the same at all ages.

Variations in the range of values at different ages may be due to basic psychological differences in the understanding of the words (or phrases) or in the interpretation of the particular ideas. If half the seven-year-old group did not understand what the words meant and if the other half of the group gave values similar to those given by the adults, we should find a narrower range in their average values. No doubt some of the seven-year-olds did not fully understand what was required of them, but we do not think that the comparatively narrow range of values at this or any age is due to such a factor of misunderstanding. One bright youngster, aged six and a half years, studied individually, certainly understood what was wanted and was quite familiar with the words. The ratios for 'nearly always' to 'very rarely' based on his answers are 1·7 for Raining; 14 for Cold; and 2 for Late. Two of these ratios fall within the appropriate range for his age group, and are quite different from those obtained from any adult. We should add that where it was perfectly clear to us that the children did not understand what they had to do their answer sheets were excluded from the analysis. We could easily tell whether their mistakes were of the kind which showed that they did not understand the questions. No child's replies were excluded because of the magnitude of the figure he gave.

Of the three categories studied, the one which shows most markedly a growing understanding with age, from seven to eleven years, is Frequency. This may be due to the fact that the interpretation of this category depends on seeing a relationship between an event and a period of time, not, as in the other categories, on the grasp of a single idea. Indeed, Frequency may be said to involve Quantity and Duration. At the age of seven, the child does not distinguish between 'nearly always' and 'very rarely' or 'seldom'. The ratio of the values given by children, on the average, to these two extremes approaches the value 'one'. At the age of ten, the value for 'nearly always' is, on the average, three or four times as large as 'very rarely' or 'seldom'. Between seven and ten there is a steady tendency in this direction, but there is a long way to go before the adult ratio of 20 is reached.

Table 9. Category: Frequency. Ratio $\frac{\textit{nearly always}}{\textit{very rarely}}$ *based on values given by children and adults.*

Context	Children, average age				Adults
	7·1	8·4	9·4	10·1	
Raining	1·3	2·6	3·2	3·2	19·6
Late[1]	0·9	2·8	3·6	4·7	20·0
Cold	1·5	3·0	3·0	3·3	20·4
Average	1·2	2·8	3·3	3·7	20·0

[1] 'Seldom' substituted for 'very rarely'.

The effect of increase in age is shown more in relation to the understanding of 'late for school' than of 'having a cold' or 'raining'.

In the expressions for Duration, the most striking feature is the enormously wide range of absolute values given by individual children. The ratio of the two extreme expressions ('in a moment' and 'for a long time') is very large, and increasingly large with the older children. The ratios vary with context.

If we plot the values of the ratio $\frac{\text{nearly always}}{\text{very rarely}}$ against age, for children and adults, we find that they more or less fall on a straight line, though any irregularities during adolescence are not here considered. (See Fig. 5.)

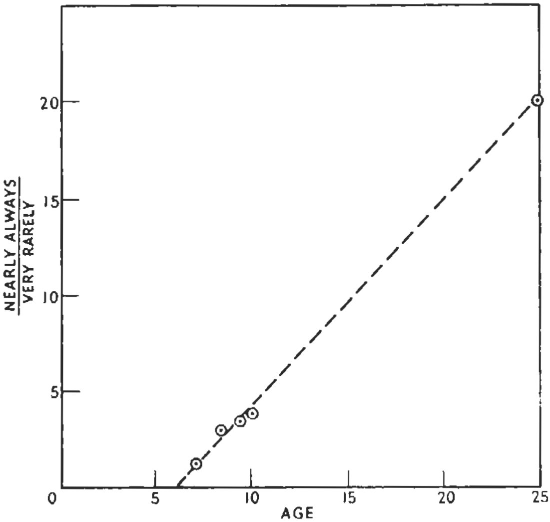

FIG. 5. Frequency: ratio $\frac{\text{nearly always}}{\text{very rarely}}$ plotted against average age

If we look at the foregoing studies in a more general way we can see that they open up possibilities of investigating the use of language in new ways. One such possibility is an experimental semantics, which would study the meanings we attach to words by the way we actually 'operate' them. At the same time one could study the 'operational' use of words at different ages during the period of development. Take the following example. In our experiments we have asked children or adults to write down a number in place of a word or phrase. We could instead present them with a pile of marbles and ask them to pick out 'some', or 'many', or 'a lot'. We could use other kinds of content and with populations of content varying in size. We could ask a child to take some marbles from a group of 5, 50, or 500 marbles. The operational use of words could also be studied when a person is alone or when others are present. A child, when alone, may take a certain number of marbles when he is told to take 'some', but he may take less if there is another little boy waiting to take his 'some'.

This method of investigation is applicable not only to such categories as we have studied, but to a much wider range of expressions. Our preliminary results suggest that many of these terms are not used as absolutes but rather in the sense of a proportion of a total number. When a child uses the expression 'a lot of' he might have in mind a ratio of something like three-quarters, whatever the total amount involved, and when he is talking about marbles or about the boys in his school. This hypothesis could be submitted to closer test by the experiments we have suggested.

The use of words and phrases in a relative rather than an absolute sense resembles a tendency characteristic of many psychological phenomena. It seems to be a basic feature of perception to perceive relationships rather than absolutes, for example, in judging magnitudes, brightness, or heaviness. In order for the subject to give equal differences in his responses, the stimulus has to be increased geometrically. Two aeroplanes flying overhead do not sound twice as loud as one, nor do three sound three times as loud. In order for the sound of the aeroplanes to seem to the listener to increase by equal steps, the number of aeroplanes flying overhead would have to be increased in some kind of geometrical proportion. When we describe something and appear to be using an adjective in an absolute sense we seem to have an implicit standard in mind. If we say that a woman is fat, we have in the back of our minds an implicit idea of an average woman, neither fat nor thin, and this model forms the basis for our description of this lady as fat. The implicit norms vary from person to person. A *very* fat lady might describe a fat one as thin and an *exceedingly* thin lady might describe a fairly slender one as fat. The words young and old change in their subjective meanings with age because of a moving standard.

When a child learns the meanings of the kind of word under consideration, his task is different from learning words which have a highly precise meaning, like words referring to definite objects such as tables or chairs or words which refer to numbers. The word 'two' or 'three' has to be learnt in a different fashion from a word like 'some', the meaning of which has to be derived and generalized from a variety of usages in different situations; and the child does indeed learn to use these vague terms in a diversity of contexts.

It is also possible to examine differences in the usages of words in different social or economic classes. When a poor child says "I have some toys", he may have a different number in mind than a rich child who makes the same remark. We could perhaps discover the implicit norms underlying these usages in different social classes. Such methods of study point the way to the possible measurement of individual and social norms of judgment and of the interrelationships between them.

Chapter Five

INTENSITY OF BELIEF

We all hold a large variety of beliefs. We have beliefs about number, time and space, about things that are supposed to have happened in the past, about places that are supposed to exist a long way off, about the goodness or badness of different actions, about the beauty or ugliness of different things, about our own capacities, about the shortcomings of others, and many other kinds of belief. Each belief is held with a particular degree of conviction. A person who accepts without question the statement: "two and two are four", as correct may have some doubt about the statement: "Noah built the Ark" and even more misgivings about the statement: "all politicians selflessly devote themselves to the public service". The same statement might be accepted wholeheartedly by some, rejected outright by others, and accepted with a grain of salt by a third group.

Consider the following statements:

(i) Queen Anne lived
(ii) Anthony loved Cleopatra
(iii) There is an island called Borneo in the Far East
(iv) Bacon wrote Shakespeare
(v) Television is a great social evil
(vi) Tobacco smoking is the main cause of lung cancer.

A person may describe any of these statements as 'true' or 'false' or as 'possibly' or 'probably' true. He might even be able to give a fairly precise indication what he means when he says 'possibly true' or 'probably true'.

Let us take another example. A housewife is about to go shopping to buy ten different articles at ten different shops. Before she embarks on her expedition, we ask her: "How many of the ten shops do you think will have in stock the goods you want to buy?" Suppose that she replies: "Seven of the shops." We then ask: "How sure are you of this?" Her reply to this question might be: "I'm quite certain" or "I am rather doubtful" or perhaps "I am very doubtful." If she replied that she was quite sure, we could go on to ask: "Are you sure that you are sure?", and then: "Are you sure that you are sure that you are sure?", and so on; and if she said "I am very doubtful", we could ask her: "Are you sure that you are very doubtful?" and: "Are you sure that you are sure that you are doubtful?"

These examples are intended to suggest that we cannot make any choice, judgment or decision without at the same time having some degree of confidence in it. The degree of confidence may range from absolute conviction in its truth or in its falsehood or it may have some measure of uncertainty. And the degree of confidence is itself held with a second degree of confidence which may be the same, higher or lower. Similarly, this second degree of confidence is in turn held with a third degree of confidence at another level, and so on. Any statement of belief should therefore be supported by information which indicates the strength of conviction with which it is held.

A high level of confidence at the second stage may sometimes turn out to be a handicap. We are reminded of a peer accused of murder who conducted his own defence in which he endeavoured to prove that he was insane. He was anxious to impress this on the jury, and his efforts were so successful that they led to an outcome which he did not anticipate. The jury decided he must be sane otherwise he could not have made such a brilliant demonstration of his insanity.

Although we are here primarily concerned with intellectual states of mind, we are not suggesting that such states are actually separable from emotional states. It is only for convenience of discussion that intellectual aspects have been delimited. We must, however, make a brief reference to some basic features of doubt and uncertainty.

Very roughly speaking, people may be divided into three categories: the dogmatists, the sceptics, and the doubters. The dogmatist confidently asserts, the sceptic with equal assurance denies, and the doubter—often no less confidently—doubts.

These intellectual forms appear to reflect underlying emotional patterns. Perhaps the doubters are best understood because it is they who tend to come for psychiatric help. Intellectual doubt, when it is constant and deep-seated and irrespective of subject-matter, seems to be connected with an unresolved emotional conflict. The doubter may be, and often is, unaware of the emotional roots of his doubt. This is exemplified best in obsessive persons. The obsessive carries a regular burden of doubt and uncertainty. He broods endlessly on problems unsolved and indeed incapable of solution. Frequently his doubts are rooted in unresolved sexual difficulties, as the following example illustrates:

A patient looking at a door was compelled to spend much time brooding about the problem: what is the main thing, the empty space, filled out by the door, or the substantial door, filling out the empty space? This 'philosophical' problem covered the other doubt: what is the main thing in sexuality, woman or man? And this again meant: what is the main thing in me, femininity or masculinity?[1]

When an obsessive is in a state of extreme doubt it may be that he fears the omnipotence of his own beliefs. If he admitted the belief to himself, it would, he thinks, cause reality to conform to it. He doubts the truth of a rumour of someone's death because if he believed the rumour it would (he imagines) *cause* the person's death. He is anxious to prevent himself having thoughts which might bring about death. So he declares the rumour to be false. The doubt may torment him so much that when the news of death is confirmed, he feels an immense relief for now he need no longer fear that *his* thought caused the death.

The obsessive person is apt to treat any simple situation as a major crisis. He suffers agonies of indecision and enters into endless rehearsals of all that might be said for or against a possible course of action; he is well-skilled in posing false antitheses. Such considerations led Pierre Janet to wonder whether philosophical speculation itself is more than a sickness of the mind, and Freud also compared schizophrenic thinking with philosophizing.

Measuring Degrees of Conviction

The aim of the experiments we are now to describe is to examine the nature of uncertainty in relation to various levels of confidence. Before giving the experimental details let us

[1] Fenichel, O., *The Psychoanalytic Theory of Neuroses*, p. 297. New York: Norton, 1945.

consider a hypothetical instance in which some of the problems arise. Suppose we ask a friend to guess the exact number of men and the exact number of women in a vast crowd. After he has made his two guesses let us ask him to say how many of the two guesses he thinks he made correctly. The possible answers to *this* question are 0 (no guess correct), 1 (one guess correct), 2 (both guesses correct). Because the task is so extraordinarily difficult we should expect him to answer 0 (no guess correct). If, however, the task had been ridiculously easy, we should have expected the answer 2 (both guesses correct). If our friend had felt the task to be neither too easy nor too hard, we should have expected the answer 1 (one guess correct).

Let us now suppose he has been asked to guess the exact number of men and women in *another* vast crowd, thus giving us two more estimates. We can again ask him how many of *these* two he thinks are correct and again he will presumably say '0'. Here, also, if the task had been very easy he would have replied '2' correct; and if moderately difficult, '1' correct.

Our friend has given us his guesses of the number of men and women in each of two vast crowds. And he has added how many of the two guesses in each case he thinks are correct. Because of the great difficulty of the task we assume he has said that neither of his two guesses was correct in the case of each crowd: that, in other words, he had replied '0' in answer to both our second questions.

We now proceed to ask him how many of *these two '0's* will be found correct when his answers are checked against the facts. Here again the possible answers are 0, 2 and 1. Because the original tasks were impossibly hard, he is almost certain to say '2', that is, that both his present answers will be found to be correct. In other words, he is certain that none of his *first* estimates was correct.

If the two original estimates had been very easy ones to make instead of very hard, we should also have expected him to answer '2' to the later questions. If, however, the original guesses had not been very hard ones to make, then, in replying to our question: "How many of the two guesses you think are correct?" he might well have said: '1' (one correct). And if we had repeated the question in relation to two other not very hard guesses, he might again have said '1'. Finally, if we should ask him how many of the two answers '1' he thinks are correct, he might be expected once more to answer '1'.

The general problem may thus be approached experimentally by asking the subject to carry out two tasks similar in kind and difficulty. When he has finished these tasks we ask him how many of the two he thinks are correct. His answer gives us a measure of his first level of subjective probability, which we call P_1. We repeat this procedure with two more tasks similar to the first two, and obtain a second P_1. We can now proceed to the next stage and ask our subject how many of his two P_1's will be found correct when his answers are checked. His reply to this last question will give us a second level of subjective probability, which we call P_2.

An actual example of a boy's estimates of the length of a number of lines, together with his successive degrees of confidence might make the analysis clearer.

The figures under P_1 are the boy's answers to the question how many of each pair of his estimates he thinks are correct. The figures under P_2 are his answers to the question how many of the answers under P_1 he thinks are correct, or, in other

Length of line to be estimated (inches)	Boy's estimate (to nearest inch)	His subjective probabilities P_1	P_2	P_3
5	6	1	2	1
11	12			
9	8	2		
3	3			
7	6	1	1	
2	2			
8	10	1		
5	5			

words, how much confidence he has in the accuracy of his original estimates of length. The figures under P_3 have the same relation to those under P_2 as those under P_2 have to those under P_1.

In order to express these values of P_1, P_2 and P_3 on a scale extending from 0 to 1, each figure must be divided by 2 because in each case the maximum figure that can be put down as an answer is 2. The figure 2 here means complete confidence in the answer at the previous stage.

The values which the boy actually assigns and the corresponding values on the new scale are thus:

Subjective probabilities as expressed

By the boy	*On a scale from 0 to 1* (P_1, P_2, P_3, etc.)
0	0·0
1	0·5
2	1·0

The first level of subjective probability (= P_1) need not be based only on two estimates. It can be based on any number of initial estimates. We can ask the subject: "How many of your *five* estimates (or any other number of estimates) do you think are correct?" We should then have a more sensitive measure of P_1. And, in the same way, the second level of subjective probability (P_2) may be based on any number of values of P_1.

In Appendix I to this chapter the reader will find a Table setting out a hypothetical example of P_1 and P_2 arising in practice when the tasks range widely in their degree of difficulty.

The initial tasks in the experiments are of two basic types:

Type I (a). The first type includes such tasks as estimating the length of lines or the weight of objects. The subject estimates to the degree of precision demanded by the experiments, for instance, to the nearest inch or half inch, as the case may be. In this manner the ease or difficulty of the task may be varied. The subject is told that his estimate will be marked as correct or incorrect according to whether it falls within the required range of accuracy.

Type I (b). This includes tasks such as estimating which of two lines is the longer or which of two lines of poetry is the better, according to some specified standard, such as the teacher's opinion.

Type II. The second type consists of tasks in which the odds of success are fixed beforehand by the experimenter. For instance we present a card on which there are, say, the five letters A, B, C, D, E; we tell the subject that one of these letters is also printed on the back of the card. He then has to guess which of the five letters it is, and he knows that he has a one in five chance of being right. Or we say that four of the five letters are printed on the back of the card, and that he has to guess *one* of these. In this way the initial probability of success is varied.

An essential difference between these two types of task is that in the first the subject's success depends on his skill. In the second, skill plays no part in the initial guess, though some might argue that clairvoyance or telepathy may have some influence!

The differences between Type I (a) and I (b), on the one hand, and between Type I and Type II, on the other, are thus solely differences in the way the initial judgment is made. The subsequent formation of P_1, P_2 etc. is the same throughout Types I and II. The procedures followed in these different types of experiment may now be briefly described.

Instructions : Experiment Type I (a)

I am going to show some lines of different shape and length. First you will be shown two examples. You have to estimate the length of each line to the nearest inch. Make your estimate on seeing each line. I have measured the lines, so I will mark your estimate afterwards as *correct* or *incorrect.*

In one form of this experiment the task was made harder, the children having to estimate the length of lines to the nearest quarter of an inch. Sometimes common objects were substituted for lines and the children, in one experiment, estimated the weights to the nearest ounce, and, in the other, to the nearest half ounce.

In Experiment Type I (b), the children were shown in turn pairs of lines of unequal length and they judged which one was the longer. Here again they were told that the length of the lines was known and that therefore their answers would be marked. In a variation of the experiment, the children were shown pairs of lines of poetry and then judged which one of each pair was the better. They were told that their form-master's opinion would be taken as the criterion of correct judgment.

Procedure: Experiment Type II

> We show the children a series of cards in turn. On each card there are the letters A B C D E. The children are told that on the other side of the card there is one of these five letters. They have to guess which letter it is on each separate card. There were eight cards.

In one variation of this experiment the children were told that there were four of the five letters on the back of each card and they had to guess any one of the four for each card. In another variation the cards had two letters A, B, and the subjects were told that one of these letters was also on the other side of the card and they had to guess which letter it was for each card.

The procedures described above only refer to the first judgment in each experiment. When they had made two initial judgments, the children wrote down how many of these two they thought were correct, thus giving their first value of P_1. They then proceeded to make another two initial judgments followed by a second P_1. Following this, they recorded how many of the two values of P_1 they thought were correct, thus yielding a first value of P_2, and so on until a second value of P_2 was obtained, when they proceeded to give a value of P_3.

In other experimental arrangements, the value of P_1 was based not on a set of two initial judgments but on a set of three or five.

The detailed numerical results are tabulated in Appendix 2 to this chapter and the reader who would like to study the differences in intensity of belief due to kind or degree of difficulty of task may care to reflect on the figures given there. Here we shall simply interpret these results in general terms.

In the eight experiments of Type I, the groups taken as a whole said their answers would be correct roughly on about 70 per cent. of the occasions. In two of the experiments the task was deliberately made hard, the children having to estimate length to the nearest one quarter of an inch, instead of to the nearest inch; and weight to the nearest half ounce instead of to the nearest ounce. In both of these experiments the groups thought they would be correct only in about 55 per cent. of the answers.

In the Type II experiments, on the guessing of letters, the tasks were of three degrees of difficulty, easy, moderately hard, and hard. We find that as the task gets harder, confidence in the accuracy of one's performance gets less. By 'confidence' we mean, of course, the proportion of times a person thinks his answer is correct. When the task was relatively easy—so that, on the average, eight out of ten answers were bound to be right—the group as a whole said they were correct on 72 per cent. of the occasions. The corresponding figure for the moderately hard task is 59 per cent. and for the hard task 38 per cent. This is the picture we obtain from the analysis of the judgments of the thirteen-year-old Grammar School boys. By comparison, Primary School children aged ten to twelve years old give less clear results. The degree of difficulty of the task seems to have less influence on *their* confidence in the accuracy of their performance. Perhaps some of the younger children failed to grasp the nature of the task.

They may not have realized that they would be more often correct by sheer chance when four out of five letters were on the back of the card than when only one out of five was on the back of the card. Nor, perhaps, did some of them realize that they would be more often correct by chance, when one of *two* letters was on the back of the card, than when one out of *five* letters was on the back of the card.

Turning to the relationship between the first and second levels of confidence we find one noteworthy feature, though no very marked tendencies. Possibly P_2 and P_3 only assume importance near limits of absolute certainty. When the thirteen-year-old boys assert on two occasions that they have done one of two tasks correctly, they then tend to say that only one of these assertions is correct. This may be interpreted to mean that those who are uncertain at the first level of confidence are more or less equally uncertain at the second level.

Those, however, who say on two occasions that they did both tasks correctly or neither task correctly show no definite tendency, at the second level of confidence, to be certain or uncertain of accuracy or inaccuracy at the first level of confidence. There seems to be at this age a relationship between the first and second levels of confidence only when an uncertain state of mind is expressed.

Analysis of Meaning Attached to the Word 'Probability'

The above experiments have been concerned with rather unusual situations. In everyday life people do not normally attach to their assertions values of 1 or 2 or 0·5 or any other numerical index of subjective probability. They prefer to say 'such a thing is likely to happen', or 'it might happen'. Perhaps the most typical of all the wide range of expressions of this kind is that something 'will probably happen'.

We have therefore tried to discover what meaning children of the same age as those in the experimental groups attach to the meaning of the word 'probably' in a context commonly used in everyday life, namely, "the judge says the prisoner is probably guilty".

We should distinguish two possible meanings of the word 'probably'. In the present context the word 'probably' is an expression of a value of P_1, that is to say, it expresses the judge's confidence in his judgment of 'guilty'. The same is true of jurymen. When the members of a jury are discussing the circumstances of a particular case they are very rarely confronted with such clear issues that every one of them is absolutely certain either of the innocence or of the guilt of the defendant. In most cases they have to assess the relevant circumstances and attach a certain weight to each circumstance. All these weights in the aggregate yield in the mind of each juryman a judgment such that the prisoner is probably guilty. So it seems that when the juryman thinks that the prisoner is probably guilty he is, in effect, attaching a level of confidence embodied in these weights to his judgments about each of the considerations. That is, he says to himself "This consideration points to guilt", "This consideration points to innocence". In the result he has to say 'guilty' or 'innocent'. His final decision is based upon his P_1 value or first level of confidence in his judgments. In other contexts, such as, for example, "20 per cent. of this class of children will probably pass a scholarship because roughly that proportion has passed scholarships in the past", we are using 'probably' as an initial estimate, not as an expression of confidence in such an estimate.

A group of 56 subjects wrote down their interpretation of the statement, "The judge says the prisoner is probably guilty." An examination of their interpretations suggests two ways of classifying them. First, the interpretations can be classified according to the degree of certainty or uncertainty embodied in the judge's statement: and second, they can be classified according to whether the statement refers to a state of mind of the judge or to the objective circumstances or situation of the prisoner regardless of the state of mind of the judge.

Taking first the classification in terms of a continuum of grades or degrees of certainty, we can distinguish in the interpretations four main points along this continuum, though there are intermediate points as well, and it is sometimes hard to decide where along the scale a subject's interpretation actually lies. The four points are as follows. At one extreme 'probably' means complete certainty. At the other extreme it means a fifty-fifty chance of being guilty or innocent. In between these extremes we have one interpretation to the effect that 'probably' means almost but not quite certain, and another interpretation which means that the prisoner is more likely than not to be guilty. Twenty-nine out of the 56 boys (aged 13+) interpreted 'probably' to mean 'almost but not quite certain'. About half that number interpreted it to mean 'rather more likely than not to be guilty', half of that number again in terms of a 'fifty-fifty chance', and half of that number in turn as 'complete certainty'.

Number of interpretations of 'probably' guilty in terms of degree of certainty.

$N = 56$ boys aged 13 +

(i)	Certain	4
(ii)	Almost but not quite certain	29
(iii)	More likely guilty than not guilty	15
(iv)	Could be equally guilty or not guilty	8
		56

EXAMPLES OF DIFFERENT INTERPRETATIONS IN TERMS OF DEGREE OF CERTAINTY

Certain

"The judge is certain that the person who is being tried is guilty."

"The judge thinks himself that the prisoner is guilty."

Almost but not quite certain

"He is almost sure that the person is guilty not innocent."

"I think the prisoner is probably guilty means that he thinks the prisoner in about 99 cases out of 100 would be guilty."

More likely guilty than not guilty

"I think it means that the judge thinks that although the prisoner is not proved to be guilty, he is more likely to be guilty than to be innocent."

"He thinks the prisoner is more likely to be guilty than innocent."

Could be equally guilty or not guilty

"The chances are 50–50 as one can not tell for certain that the man has committed an offence or not."

"It means that it is a half and half chance of being guilty or not guilty."

Passing on to the classification in terms of subjective and objective interpretations, we find that 41 out of 56 gave subjective interpretations, 11 gave objective and four gave interpretations which contain subjective and objective elements.

EXAMPLES OF DIFFERENT INTERPRETATIONS IN TERMS OF SUBJECTIVE OR OBJECTIVE ORIENTATION

Subjective

"When a judge says probably guilty he is telling a person that he is not quite certain whether the man is guilty or not guilty."

"The judge thinks that the prisoner is guilty but he cannot be sure about it."

Objective

"The statement means that the prisoner might be the person who has committed the crime or the offence which the court is hearing."

"This phrase means that there is a strong chance that the prisoner has committed the crime."

Subjective-objective

"He means that it is quite likely that he is guilty, but he may not be guilty. He thinks that there is more chance of him being guilty than not guilty."

Note

There is a certain fundamental problem in the study of subjective probability which is too intricate to discuss in a general way in the body of this chapter. We have therefore dealt with it in Appendix 3.

Appendix I

Table 10. A hypothetical example showing how P_1 and P_2 could arise in practice.

No. of task	Degree of difficulty	Assumed performance	First level of confidence (P_1)[1]	Second level of confidence (P_2)[2]
col. (i)	col. (ii)	col. (iii)	col. (iv)	col. (v)
1 2	Very hard Very hard	both incorrect	0	2
3 4	Moderate Moderate	only one correct	1	1
5 6	Very easy Very easy	both correct	2	2

[1] Based on answers to the question "How many of each pair of answers do you think you are correct?"

[2] Suppose that there were two pairs of very hard tasks. There would then be two values of 0 in col. (iv). Similarly, if there were two pairs of moderate and two pairs of very easy tasks, there would be two values of 1 and two values of 2 in col. (iv.) The values in col. (v), for the second level of confidence, are obtained in answer to the question: "How many of the two P_1 values do you think are correct?"

Appendix 2

Table 11. Type I. Initial task : estimation or comparison, involving skill. Mean values of group.

Experiment	Nature of initial task	No. in group	Age	Sex	Subjective probabilities		
					P_1	P_2	P_3
1	estimating length of line	64	12+	Male	0·62	0·61	0·60
2	Do.[1]	62	12+	Do.	0·53	0·58	0·50
3	Do.	95	11–13	Do.	0·64	0·76	—
4	comparing length of two lines	29	11+	Do.	0·80	0·81	0·80
5	estimating weight	83	10+	Mixed	0·62	0·76	0·77
6	Do.[1]	84	10+	Do.	0·57	0·65	0·70
7	comparing lines of poetry	167	10+	Do.	0·74	0·79	0·81
8	Do.	29	11+	Male	0·77	0·82	0·73

[1] Hard task.

Table 12. Type II. Initial task : guessing whether a given letter is at the back of a card when odds of success are known.
Mean values of group.

	Degree of difficulty	Initial probability	P_1	P_2	P_3
N = 52 13+ Boys	easy moderate hard	0·8 0·5 0·2	0·72 0·59 0·38	0·74 0·66 0·56	0·70 0·65 0·59
N = 73 10–12 Mixed sexes	easy moderate hard	0·8 0·5 0·2	0·65 0·59 0·54	0·75 0·75 0·71	0·70 0·65 0·69

Appendix 3

Addition and Subtraction of Subjective Probabilities

We have not yet touched upon the question of the relationships between the subjective probabilities of different outcomes of the same event. If we represent the chance of winning a raffle or lottery by p, then $1-p$ is the chance of not winning, $p + (1-p) = 1$. If there are ten tickets, one of which has the winning number, the chance of drawing this ticket is 1/10th, just as is the chance of drawing any of the other nine tickets. The chance of not drawing the winning ticket is nine times 1/10th = 9/10. $1/10 + 9/10 = 1$. '1' thus represents the chance of drawing a ticket which may be winning or non-winning. In any situation where there are many possible outcomes, only one of which can actually happen, the probabilities of the separate outcomes add up to '1'.

Do these relationships hold in subjective probability? In general, do we think our chances of not winning get smaller as our chances of winning get larger, or *vice versa*? Or do we think about the chance of winning and losing in quite different and independent ways?

We can attempt to answer this question by examining the results of our experiment. We present the following problem:

> Three boys, George, John and Jack are going to run a race in a school sports. George has won his race five times on previous occasions, John has won two races in the past, Jack has never won a race.
>
> *Question A :*
>
> Write down the names of the boys you think will come first, second and third in this race. There can be no dead heat.
>
> *Question B :*
>
> You have put down the name of the boy you think will win the race. If you are absolutely certain that he will win, write down the number 10. If you are completely uncertain whether he will win or lose write down the number 0. Or write down any other number which indicates how sure or unsure you are.
>
> *Question C :*
>
> Now look at the name of the boy you think will come second. If you are quite certain he will come second write down the number 10. If you are completely uncertain whether he will come second or in another position put 0. Or write down any other number which indicates how sure or unsure you are.
>
> *Question D :*
>
> (The subjects were then told to do the same for the boy they thought would come third.)

In another experiment with a comparable group of subjects the following two questions were substituted for Questions C and D above:

Question E:

Now look at the name of the boy you think will come second. If you are quite sure he will *NOT* win, put the number 10. If you are completely uncertain whether he will win or come second or third put 0. Or write down any other number which indicates how sure or unsure you are.

Question F:

(The subjects were then told to do the same for the boy they thought would come third).

Thirty-two boys aged 13+ answered questions A, B, C, D, and 31 boys aged 12+ answered questions A, B, E, F.

We shall first show in Table 13 how they decided to place the winner in the race.

Table 13. Number of subjects placing George, John or Jack as winner.

Winner	Age Groups		Total
	13+	12+	
George	15	16	31
John	12	13	25
Jack	5	2	7
Total	32	31	63

There is no significant different between the judgments given by the two age groups. The values given for George, John and Jack respectively as winning are significantly different from values that would arise by random allotment. The number of previous wins influences the judgments to the extent that half of the group believe that George will win. The other half of the group, who believe that John or Jack will win, appear to be favouring the non-preponderant outcome, as explained in Chapter 2.

The next question to consider is the degree of certainty attached to the boy who is chosen to win. We find that there is no significant difference between the values given for George, John and Jack when chosen to come first in the race. The average value assigned is about 6 out of 10. Since the scale 0 to 10 used in the question represents 0·5 to 1·0 on a probability scale, the values given (S) can be converted to probability values (p) by means of the formula $p = \frac{S+10}{20}$. Linearity of the scale is assumed.

In Question E we wish to know how certain the subjects are that the boy they think will come second will *NOT* win the race. If we call the subjective probability of the second boy not winning Q, then the subjective probability of his winning is 1–Q. At least this would follow if subjective probabilities obeyed the same rules as mathematical probabilities. And similarly for the boy who is thought to come third in the race. If so, we can derive from our subjects' answers their subjective probabilities that the boys they have actually chosen as second and third will win the race.

There are three possible winners. One possible winner has been assigned a direct subjective probability of winning.

For the other two possible winners we can calculate indirect subjective probabilities of winning. The question therefore is: do these three subjective probabilities add up to '1'? They do, within the limits of random variation. So far, then, as our evidence takes us in this experiment and within the age-range studied, we have no reason to suppose that the additive character of subjective probability differs from that of mathematical probability.

Appendix 4

Pairs of Lines of Poetry used in Experiment 7

1a. Thou wast not born for death, immortal bird.
1b. The fair breeze blew, the white foam flew.

2a. He hath awakened from the dream of life.
2b. Through caverns measureless to man.

3a. There is a budding morrow in midnight.
3b. Peopled with unimaginable shapes.

4a. 'Twas pitiful, 'twas wondrous pitiful.
4b. There would have been a time for such a word.

5a. And the silken sad uncertain nestling of each purple curtain.
5b. Out of the cradle endlessly rocking.

6a. There was silence deep as death.
6b. The river glideth at his own sweet will.

7a. A long, forlorn, uncomfortable way.
7b. But who shall mend the clay of men, the stolen breath restore.

8a. Ne'er saw I, never felt, a calm so deep.
8b. Between two worlds, life hovers like a star.

9a. The precious porcelain of human clay.
9b. Man is a restless thing, still vain and wild.

10a. Every cradle asks us "Whence"? and every coffin "Whither"?
10b. Ill news hath wings and with the wind does go.

11a. A reeling road, a rolling road that rambles round the shire.
11b. The wrinkled sea beneath him crawls.

12a. Most musical of mourners, weep anew.
12b. There is a budding morrow in midnight.

13a. And after him the parson ran, the sexton and the squire.
13b. That night we went to Birmingham by way of Beachy Head.

14a. You hail from dream-land, Dragon-Fly?
14b. On the gentle Severn's sedgy bank.

15a. In sunsets golden and crimson dyes.
15b. Hope springs eternal in the human breast.

16a. When I behold the night's starred face.
16b. Forlorn, the very word is like a bell.

Chapter Six

RISK-TAKING

A circus audience breathlessly watching a tight-rope walker believes he is taking a dreadful risk. The performer himself takes his act nonchalantly, and the audience is far less sure than he is that he will succeed.

To say "there is a risk" means we are in a state of uncertainty. We are not sure what the outcome will be. The degree of uncertainty may vary widely from one situation to another and from one individual to another. To the extent that there is some uncertainty of success in anything we undertake, there is an element of risk present. Just as we can compare subjective with mathematical probabilities (see Chapter Two), so we can compare uncertainty of success in performance with what actually happens. Uncertainty of success in performance is subjective probability embodied in action. What *actually* happens gives us a measure of objective probability. When we say "what actually happens" we mean either an individual's average success in performance or the average success in performance of people in general. Life insurance systems are based on this second form of objective probability.

Uncertainty of success may be influenced by various factors such as familiarity with the task, tolerance of uncertainty, 'difficulty' of the task and awareness of the consequences. A person may be totally unfamiliar with the task in hand (or similar tasks) or he may have a mastery of it or some degree of familiarity with it. People vary as well in their tolerance of uncertainty. One person—the one who feels a need for assurance—has to be very certain of whatever action he takes and can only act when he feels highly confident. Another person never feels certain of success or failure and tends to place a moderate confidence in his activities. The appeal of fiction rests largely on the delight which people take in the thrill of suspense. The tight-rope walker in the circus produces such thrills in the onlooker. So do the serial films which end each week with the heroine perilously suspended at the edge of a precipice and about to fall into the gaping jaws of crocodiles below. Some people intensely dislike seeing the end of a film before the beginning and the middle. If they enter in the middle of a performance they keep their eyes closed until it is over and open them when the film starts from the beginning. Some people equally dislike being told the plot of a film or book which they intend to see or read.

When we say that uncertainty of success may vary with the difficulty of the task, we may measure 'difficulty' *either* in terms of a person's actual performance *or* by some absolute standard in physical units, such as, for instance, the height of a bar to be jumped. The first method measures subjective difficulty and the second, objective difficulty.

Some people are acutely sensible of possible disastrous consequences or dazzling success. If a man can clearly visualize himself as a blood-smeared corpse under a lorry, he will perhaps be a little circumspect before he crosses a road. Many curiosities of behaviour arise in this sort of way. A man may refuse to touch a door-knob unless he uses a handkerchief to protect his hand from infection. Another person possessed by dreams of vast riches may underrate the chances of being caught and overrate his chances of success in a gamble. Prison cells house some dreamers who have embezzled their employers' funds in order to bet on a horse.

We should explain what we mean by 'risk-taking' in relation to subjective probability. In the study of subjective probability one of our tasks is to determine the different mental structures which appear at different ages during the period of development. These structures affect our subjective probability values in specific situations when we are to make an estimate, a prediction, guess or judgment. (See Chapter Two.) In the study of risk-taking we are concerned with the way these mental structures are realized in our actions. It is the relationship of subjective and objective probabilities which gives rise to concepts of risk-taking.

Experiments on Risk-taking

Perhaps it will make the nature of the experiments a little clearer if a hypothetical situation is described which exemplifies some of their features. Imagine a man with his car in a large open space in which there are two poles half a mile apart. We ask him whether he can drive his car at 5 m.p.h. between the two poles without touching them. He answers "Yes". We now set the two poles a foot apart and repeat the question to which, this time, let us suppose, he replies definitely: "No".

The distance between the poles is now increased bit by bit, and each time he is asked the same question. Soon a range of distances is found at which he no longer gives a definite "Yes" or "No", but a reply such as "I may be able to", "I think I could", "Perhaps" and so on. When he gives uncertain answers of this kind we can go on to ask him how many times out of ten he thinks he could succeed. Some people would have spontaneously replied to the *first* question "nine times out of ten".

Let us now ask him to drive his car ten times between the poles without touching them; first, when the poles are half a mile apart; second, when they are a foot apart; and third, when they are two yards apart, and at other distances slightly greater than the width of the car. His performance under these diverse conditions may then be compared with his initial estimates.

We can easily conceive of variations which would introduce a considerable element of danger. If the man had to drive a high-powered car at 60 m.p.h. between two huge boulders fairly close together, his estimates of performance would doubtless differ from those which he would make if he had to drive the same car at the same speed between two marks chalked out on the road.

Experimental situations without danger can reproduce in the laboratory the circumstances of the man with the car in the field, when he is in no danger.

Ball-pushing Experiment

The subject's task was to hit a ball-bearing ($\frac{3}{4}$-in. diameter) through an opening with a mallet. The opening was at the far end of a table and the ball-bearing rolled across the table towards the aperture. The task could be made more or less difficult by making the opening narrower or wider. In fact, it was presented at 18 different sizes. The smallest size was 1 inch, and the largest, 28 inches. All subjects were quite sure that they would never fail to hit the ball-bearing through the widest opening. Before they were told exactly what they had to do, they were given five minutes in which to manipulate the mallet and ball-bearing and get used to them. But during this interval of five minutes they did not practice aiming at the aperture; they merely aimed at another ball.

Each subject had to estimate in advance how often he thought he would succeed in hitting the ball through the opening. He was asked to say how many successes he would achieve in a hundred attempts. After making all his estimates, he carried out the task at six levels of difficulty, when the opening was 2, 4, 6, 10, 16 and 24 inches in size respectively. At each size of opening the subject made five attempts. The sizes of the opening were presented in random order, not in order of increasing or decreasing size. This procedure was repeated four times, so that we had 20 attempts by each subject at each of the six levels of difficulty. Twenty-five subjects, male and female, took part in this experiment.

We can now compare what each subject thinks he is *likely* to achieve with his *actual* performance. Do people on the whole tend to behave alike in this respect? Do they all tend to think they will do *better* than they actually do, or do they think they will do *worse* than they actually do, or are they more or less realistic in their estimates of likely performance?

There are two ways in which the difficulty of the task may be measured. One way is an objective one, in terms of the width of the aperture; it is easier for anyone to hit a ball through a wide opening than through a narrow one. The other way is subjective, in terms of any person's actual performance. An opening of a given width may be easy for some people and hard for others. Such differences must be taken into account when evaluating what a person imagines he could do with what he actually achieves in practice. As we are here comparing estimates with achievement we have measured difficulty subjectively in terms of each person's *actual* performance.

An analysis of the estimates in relation to performance at different levels of difficulty suggests that there are marked individual differences. We can classify the subjects into three groups. One group, the largest (11 out of 25) give consistently lower estimates of success than is shown by their actual performance. One subject, at the level of difficulty when he succeeded in 10 per cent. of his attempts, estimated that he would succeed in only 3 per cent. of the attempts, and this same underestimate is apparent at each level of his performance, when he succeeded 20 per cent. of the times, 30 per cent. of the times, and so on, up to 90 per cent. of the times.

The second group of subjects (6 out of 25) do the exact opposite. All their estimates are greater than their actual performances. They estimate that they will succeed more frequently than they do succeed. One subject estimated that he would succeed 17·4 per cent. of the times, when he only succeeded 10 per cent. of the times. His overestimates were consistent at each level of performance.

Thirdly, there is a group of subjects (8 out of 25), who overestimate at the difficult levels, when they only succeed 10, 20 or 30 per cent. of the times, and under-estimate at the easier levels when they succeed 60, 70, 80 or 90 per cent. of the times.

If we consider the 25 subjects as a whole, we notice a general tendency. When their average performance is low because they all find the task hard and they succeed only in about 30 per cent. of their attempts, they then tend to overestimate likely success in performance. When, however, they succeed in hitting the ball through the opening more frequently than in 30 per cent. of the attempts, there is a general tendency to underestimate likely success in performance. Under the conditions of this experiment the subjects as a whole appear most realistic when they succeed about once in three times. It is at this level of success that their estimate is closest to their actual achievement.

If these subjects consistently behaved in similar fashion in a variety of experimental situations, we could infer that their behaviour was due to some basic characteristic. We could argue that the discrepancy between a person's estimates of performance and his actual performance is an important index of his sense of reality and risk-taking propensities.

Like some of our subjects, there are no doubt many people who take a poorer view of their capacities in everyday situations than is justified by the facts. They think themselves more or less inferior as compared with what they actually can do; they are modest in their claims. Others are optimistic and consistently exaggerate what they are likely to achieve. They have an inflated notion of their capacities, higher than is justified by the facts. There may be a third group of persons, represented perhaps by some of our subjects, who overrate their capacity when they feel the task is hard, and underrate it when they feel the task is easy. Perhaps there are others who, conversely, underestimate their performance when the task is easy, but such persons are not represented among our subjects.

In the experiment just discussed the subject cannot possibly come to any harm. It represents one extreme of risk-taking where danger is negligible. We were naturally anxious to investigate risk-taking at the other extreme when there is a possibility of much danger but subjects are understandably reluctant to expose themselves to situations from which they may emerge in mutilated form. It was essential, however, to have an element of real danger and this was obtained in a gymnasium. The task was to jump over a stout wooden beam (cross-section about $2\frac{1}{2} + 6$ inches) on fixed supports. The reader's imagination will help him to appreciate that failure to perform this task successfully could have exceedingly painful consequences. There was no mat or other protective device and no one to help the jumper. Before jumping, the subject had to estimate the number of times out of ten he thought he would be likely to clear the beam. The beam was moved to different heights and the subject made his estimates in relation to each height. His attempts were made at different levels of difficulty. Six different heights were presented ranging from 38 to 58 inches, in steps of four inches. The subjects gave their estimates in advance for all the different heights before attempting any height. The heights were presented in random sequence, and so the estimates and attempts were also made in random sequence. Three attempts were made by each subject at the six different heights of the beam.

The subjects consisted of 18 male undergraduates and the experiment was carried out in the University gymnasium.[1] As in the preceding experiment, our main object was to compare each subject's estimates of his likely success in clearing the beam with his actual performance and to make this comparison at each of the different levels of difficulty.

It should be made clear that we are not concerned with

[1] Located in the University Department of Physical Education.

situations so easy that subjects believe they will always succeed, and they do succeed. Nor are we interested in situations so hard that the subjects think they can never succeed and make no attempt. The heights chosen were such as to lie, for most subjects, within a range of possible success or failure.

When we inspect the subjects' estimates of their *likely* performance and compare them with their *actual* performance we find that 12 of them mostly underestimated their likely performance in jumping the beam, three mostly overestimated, two gave accurate estimates and one underestimated the three lower heights and slightly overestimated the fourth. If we regard this last individual as having in the main underestimated, there are 13 subjects who underestimated, three who overestimated and two who gave realistic estimates. In comparing the results of this dangerous experiment with the results of the safe ball-pushing experiment, we find that statistically there is a significantly greater tendency to underestimate likely performance when there *is* danger than when there is not.

If, as in the ball-pushing experiment, we now combine the individual estimates at each difficulty level into a single group estimate we find the most realistic forecast or assessment of performance when the subjects succeed in about 30 per cent. of their attempts. When they succeed less often they tend to underestimate. When a task is either very hard or very easy the subject's estimates, as we should expect, agree with his performance. As he moves from a very hard task or from a very easy task he tends to overestimate or underestimate respectively.[1]

[1] These conclusions are supported by experiments on specially designed card games in which the mathematical probabilities were

Daily Risk-taking

Life begins with a situation of binary outcome; will the ovum be fertilized or not? Thereafter hazards never cease. We shall select for study one common hazard, that of crossing a road against approaching traffic.

We measured the presence of risk, not in relationship to the proportion of successes in performance, but in terms of whether a person crossed the road or not. The only assumption we needed to make was that pedestrians, on the average, when they show signs of intending to cross a road presumably desire eventually to arrive at the other side. Pedestrians in such situations were totally unaware that they were under observation.

The site chosen for the study was a main road in a fairly busy part of Manchester. In the centre of the road there was an island but no form of traffic control, and the density of traffic at the site was moderately heavy. The study extended over a period of about a month in which 1189 observations, were made on 491 persons. Each kind of observation related to the age and sex of the pedestrian, whether he crossed or not, and the speed and distance of the vehicle. The age of the pedestrian was guessed from his appearance or attire. The speed and distance of the oncoming vehicle were judged by recording the time from the moment the pedestrian stepped to the edge of the road and looked towards the approaching traffic to the moment when the first vehicle reached the site.

known. It was found that the subjects overestimated when the chance of winning was less than 25 per cent., and underestimated when the chance of winning was more than 25 per cent. When the chance of winning was about 25 per cent. the subjects' estimates were most realistic. Preston, M. G. and Baratta, P., "An experimental study of the auction value of an uncertain outcome." *Amer. J. Psychol.*, 1948, 61, 183–93.

Preliminary study had shown that when a vehicle was more than ten seconds away pedestrians, without exception, would cross the road. We have therefore limited the present study to circumstances when vehicles were ten seconds or less away from the site. Our observations of the pedestrian's behaviour, whether he crossed or did not cross the road, were made in relation to the speed and distance of the vehicle. We ignored pedestrians crossing in groups. There were comparatively few bicycles on this road as compared with motor vehicles so we excluded from our analysis observations of pedestrian behaviour in relation to oncoming cyclists.

In Table 14 we show the number and percentage of people crossing when the vehicle was at different distances away. Some idea of the density of the traffic may also be obtained from the figures in this Table. Of 1189 instances in which a person looked up at the approaching traffic with a view to crossing the road, 205 looked up when the vehicle was about two seconds away, 184 people looked up when the vehicle was about three seconds away. By inspecting the first column of Table 14 one can immediately see how far away a vehicle was when the people looked up. We may infer from the Table that the most likely distance of the vehicle when a person looked up would be about two or three seconds.

What proportion of people who have the opportunity to cross do in fact cross when the vehicle is at different distances? Only one person out of 364 who had the opportunity of crossing when the vehicle was about two[1] seconds away did, in fact, cross. When the vehicle was four or five seconds away

[1] Strictly speaking, each time in the Table denotes an interval of one second. Thus two means between one and a half and two and a half seconds.

about half the pedestrians who had the chance of crossing did so. When the vehicle was about nine or ten seconds away almost everyone crossed.

Table 14. Number and proportion of people crossing the road in relation to the time taken by the vehicle to reach the site.

Time taken by vehicle (seconds)	Number of observations	Number crossing	Percentage crossing
1	159	0	0·0
2	205	1	0·5
3	184	22	12·0
4	159	60	37·7
5	110	65	59·1
6	94	76	80·9
7	63	58	92·1
8	40	37	92·5
9	35	34	97·1
10 or more	140	138	98·6

When a person has to cross a road, he would never try to do so if he thought there was no chance at all of getting to the other side safely, unless he were contemplating suicide. On the other hand, only very rarely do we have a sense of complete and absolute safety when we cross a road. Not that we consciously calculate the risks or reflect on our chances of reaching the further side intact, but most of us display a heightened awareness of possible danger in our circumspect steps and other manifestations of vigilance. This behaviour may be said to embody a certain level of risk-taking. The mere fact that we look to see if traffic is coming shows that we are aware of a lurking peril. It may be assumed that there is a *least* distance[1] from the vehicle at which an individual will cross the road. This least distance may vary from person to person but will presumably be more or less similar for the same person from occasion to occasion. This distance indicates what we shall call his *maximum risk-taking level*. It indicates the greatest risk he is in the habit of taking.

The reader should turn to Table 14 and consider the fourth column. This states the percentage of people crossing a road out of the total number that could have crossed, at each distance from the vehicle. Any percentage here includes the percentage that would have crossed at a *shorter* interval. For example, 59·1 per cent. crossed when the vehicle was five seconds away, but these would almost certainly have crossed when the vehicle was further away; 80·9 per cent. crossed when the vehicle was about six seconds away and we can be almost certain that these would also have crossed when the vehicle was seven, eight or more seconds away.

Let us now note that 37·7 per cent. crossed when the vehicle was about four seconds away. The difference between 37·7 and 59·1 is 21·4 per cent. This proportion of pedestrians would have crossed at five or more seconds, but they would not have crossed at less than five seconds. In other words 21·4 per cent. have their greatest level of risk-taking at about five seconds. It we take our sample as a whole, the greatest risk-taking level of 50 per cent. of them is reached when the vehicle will take about 4·6 seconds to arrive at the place where they are standing. And in our sample 50 per cent. will not cross when the vehicle is less than 4·6 seconds away.

Any person meeting a hazard in daily life probably tends to

[1] By 'distance' here we mean a function of the distance and velocity of the vehicle which we measure in terms of time.

acquire a given level of maximum risk-taking in that type of situation. A person will jump off a bus or train when it is going at less than a certain speed. He will arrive at work not later than a certain interval before work begins. These levels vary from individual to individual. The maximum risk-taking level for any person is the point at which he will begin to assume upon himself the given risk. As a rule, a person will not begin to act when the hazard is at a level greater than his maximum. As the level of risk gets less, he will tend to act all the more readily. We might expect that the performance representing maximum level of risk-taking is much more realistic when there is no danger. A person will probably walk more quickly along a plank placed on the ground than on the same plank which bridges the roofs of two high buildings.

What sort of variations can we expect in maximum risk-taking levels? Our investigation at least enables us to answer this question in relation to a common situation which everyone encounters.

Table 15. Maximum risk-taking levels

Time taken by vehicle (seconds)	Percentage crossing	Percentage crossing at not less than the stated intervals
1	0·0	0·0
2	0·5	0·5
3	12·0	11·5
4	37·7	25·7
5	59·1	21·4
6	80·9	21·8
7	92·1	11·2
8	92·5	0·4
9	97·1	4·6
10 or more	98·6	1·5

Our sample consisted of a very mixed collection of pedestrians of all ages and both sexes. We can see from Table 15 the relative frequency of maximum risk-taking levels in this group. Only one in 200 has a maximum risk-taking level of 2 seconds. Such a reckless person would only escape by the skin of his teeth. As already mentioned, 50 per cent. have a

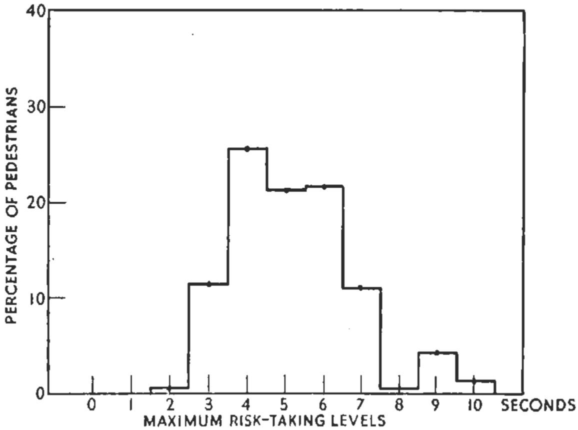

FIG. 6. Distribution of maximum risk-taking levels

maximum risk-taking level of 4·6 seconds, which allows them a margin of safety of about 2 seconds. About five or six per cent. require a level of nine or ten seconds. Among the male pedestrians we find that those aged 16 to 30 take a bigger risk than those aged 31 to 45. Among the females those taking the highest risk are in the 16 to 30 age group. Males take bigger risks than females. We should emphasize that by greater or less risk we mean that they allow themselves less or more time to cross the road.

In this chapter we have given only a few instances of how one might attempt to discover the principles underlying risk-taking. But, of course, the taking of risks pervades most human activities from birth to death. On this vast subject we can only touch briefly, first in relation to our repertoire of habits. We acquire many of our habits by profiting from our past successes or failures as, for instance, in learning to use a typewriter. But there is a class of habits which we learn almost entirely by profiting from successes, because failures would be disastrous. Some failures must be avoided at all costs because they would end in death. Not only alpine mountaineers, steeplejacks and pioneer airmen have to acquire habits in this way. Every motorist has to learn to pass and overtake cars and to come out of skids without failing.[1]

In criminal behaviour the degree of risk which a man is prepared to take may well be the decisive factor in making him commit a crime. In contrast to his law-abiding neighbour, he is the sort of person who may think he is unlikely to be caught; and that, if caught, unlikely to receive a severe sentence. One factor in the making of a habitual criminal may be the gradual rise in his maximum risk-taking level after committing a succession of undetected crimes.

There is another domain of human activity which affects not just a small proportion of the community but almost everybody. Statistics[2] of pre-marital intercourse suggest that, at any rate in the institutions of higher learning of the United States, large numbers of young persons live a sexually

[1] See Katz, David. The Psychology of the Margin of Safety, *Presidential Address, Proceedings of the Thirteenth International Congress of Psychology*, 1951, pp. 33–49.

[2] The reader who doubts the authenticity of this statement may care to consult the two Kinsey reports.

hazardous existence. Nowadays when family control is the rule, the whole of sex life is often fraught with anxieties connected with the risk of producing an unwanted pregnancy. Those who place their faith in a particular form of contraceptive might be somewhat dismayed if they saw a manufacturer's advertisement in a medical journal proudly claiming for his product an efficiency of 99·82 per cent.

The element of risk is profoundly involved in the world of play: "to dare, to take risks, to bear uncertainty, to endure tension—these are the essence of the play spirit".[1] The world of sport is distinctive in that the participants are often prepared to take great risks. Mountaineers, skiers and rugger players endanger life and limb though in private they may be as cautious as anyone else. One wonders whether an athlete like Roger Bannister ever exploited his reserves of stamina in ordinary life as he did on the track. The effort in play is unstinted, lavish and prodigal. It is not calculated with an eye to the amount of profit. In it the maximum risk-taking level is at its height. Wherever human effort is unbounded, in the realm of art or science, in sport, in saintly self-abnegation, or in obscure heroism, the elements of play and risk reach their full expression.[2]

[1] Huizinga, J., *Homo Ludens*, p. 51. London: Kegan Paul, 1943.

[2] Cohen, John, 'Ideas of Work and Play', *Brit. J. Sociol.*, 1954, 4, 312–322.

Chapter Seven

GUESSING AND ESTIMATING

We are now ready to study the types of 'guess' which are made in situations of varying complexity. Such guesses range from choices between two alternatives, at one extreme, to judgments which may involve some intuitive understanding of a frequency distribution, at the other. The reader may be reminded of a familiar example of a frequency distribution obtained by grouping the adult population into classes differing in height by one inch.

The study of distributions is closely bound up with that of sampling. Here, too, we are concerned with the question: how does knowledge about a population affect guesses about samples drawn from it? We can compare guesses made when nothing at all is known about the population with guesses made when the guesser has different kinds and amounts of information about the total population.

If nothing at all is known about the nature or composition of the population, we can infer from the guesses what assumptions are made about the population itself. Imagine a large box on a table. The experimenter tells his subjects that he has no idea about the contents of the box and asks them to guess what he is going to take out of it. It is rather unlikely that anyone will say that he is going to take out an elephant. Even when no information at all is given about a population, the range of guesses about its nature is not unlimited.

In this chapter we shall study how children of different ages guess or estimate how often certain events are likely to happen in various simple and complex situations. Suppose the experimenter has in front of him a large urn containing blue and yellow beads. Different kinds of guess or estimate will be made according to one of the following three types of condition.

Type I.

He withdraws the beads from the urn in sets of four. He places each one in a separate beaker, and the subject guesses the colour of each of the four beads. There are 16 sets of four beads.

Type II.

He withdraws the beads from the urn in sets of four and *each set of four* is placed in a separate beaker. There are 16 beakers in all. The subject now guesses the number of *blue and yellow beads* in each beaker.

Type III.

The beads are placed in 16 separate beakers, as in II above. The subject this time guesses the number of *beakers* containing 0, 1, 2, 3 and 4 blue (or yellow) beads respectively.

In these types of situation we have either told the subjects nothing at all about the proportions of blue and yellow beads in the urn *or* we have told them that the proportions are equal *or* we have told them that there are three times as many blue as yellow beads in the urn. Of course, many other variations are theoretically possible.

Let us now examine in turn the nature of the estimates made by children of different ages under these three types of condition:

Type I: The procedure takes the form of a class demonstration in which beads are drawn singly from a large urn containing blue and yellow beads, and the children guess the colour of each bead. When the children are not told the proportions of blue and yellow beads in the urn the instructions they receive are as follows:

> There are a lot of beads in this urn but we do not know exactly how many. They may be blue or yellow, but no other colour. There may be no blue beads or any number of blue beads. There may be no yellow beads or any number of yellow beads. We do not know the proportions of yellow and blue beads in the urn. We shall take four beads at a time from the urn and place each one of these in a separate beaker. Here are four beakers, one for each bead.

The children can hear the beads drop into the beakers but they cannot see the colour of the beads at any stage of the experiment. After they have recorded their guess of the colour of the bead in each of the four beakers, the beads are replaced in the urn. This procedure is repeated 16 times. In other experiments, when the children are told the proportions of blue and yellow beads in the urn, the instructions are varied accordingly.

I (a). *Proportions of blue and yellow beads stated to be unknown.* The first point to note is that when no information is given about the composition of beads in the urn, and the subject does not know what the proportions are, he seems to assume that they are equal. We have encountered this phenomenon in earlier chapters. The proportions of blue and yellow beads *actually* allotted by the children vary to some extent from child to child, but this variation is a relatively minor one. In fact, two-thirds of the proportions allotted for the blue beads in the urn lie between 46 per cent. and 58 per cent. This represents a comparatively narrow scatter.

We should bear in mind that the fact that the average proportions allotted by the children are not exactly 50 : 50 but 52 per cent. blue and 48 per cent. yellow may be due to a number of influences. Some children may prefer blue to yellow, or they may prefer to write down the letter B for blue rather than the letter Y for yellow. It may be that the wording of the instructions affected them to some extent. The combined effect of influences such as these leads to an average proportion given by all the children of 52 per cent. blue, which differs significantly from the theoretical expectation of 50 per cent.

The second point to note is the preference for certain kinds of arrangement and the avoidance of others. There are 16 possible arrangements of four beads, each one of which may be blue or yellow. If we disregard colour (which in the present study is irrelevant), there are eight possible arrangements, viz.,

(i) B B B B or Y Y Y Y } homogeneous or
B Y Y B or Y B B Y } symmetrical

(ii) B Y B Y or Y B Y B } alternating or
B B Y Y or Y Y B B } repetitive

(iii) B B B Y or Y Y Y B }
B B Y B or Y Y B Y } asymmetrical
B Y B B or Y B Y Y }
Y B B B or B Y Y Y }

When we analyse the frequencies with which the above three groups of arrangement are given by the children (*Table* 16, col. 2), we find the most marked preference for homogeneous or symmetrical arrangements. Next in order of preference is the alternating or repetitive arrangement, which is balanced without being symmetrical. The least popular choice is the third category, the asymmetrical kind of arrangement, described by some children as 'ugly' or 'unfair'. This strong tendency to choose symmetrical or alternating arrangements in guessing sets of beads drawn from an urn resembles closely the behaviour of the children when they guess on which side of the display board a light will appear. (See Chapter Two.)

An influence which may affect the preference for particular kinds of arrangement is the tendency to repeat the previous choice of blue or yellow in a set of four beads, rather than to change to the other colour. Children appear to prefer to choose the same colour after any given choice, within any set of four beads.

From the foregoing results it is evident that the choice of colour for each bead in the sets of four drawn from the urn is not made at random. The non-random nature of the guesses or choices is due to a liking for arrangements which apparently have an æsthetic appeal or which seem fair and a dislike for those that seem unæsthetic or unfair.

I (b). *Proportions of blue and yellow beads stated to be in the ratio of* 1 : 1. What happens when we tell the children that the proportions of blue and yellow beads in the urn are equal? The result is a marked increase in their choice of arrangements with two blue and two yellow beads. The number of such arrangements given by the children is significantly greater than the number that would occur by sheer chance. Each set of four single beads in separate beakers seems to be thought of as a miniature of the population of beads in the urn, so far as proportions of the two colours are concerned. This explanation is supported by the reasons afterwards given by the children in reply to questions about their choice of different arrangements.

I (c). *Proportions of blue and yellow beads stated to be in the ratio of* 3 : 1. The main effect of telling the children that there are three times as many blue as yellow beads in the urn (as compared with telling them that there is a 1 : 1 ratio) is to induce them to make a much larger number of arrangements with three blue and one yellow bead. The increase significantly exceeds the number of such a type of arrangement that would be expected if there were no special preference for a 3 : 1 ratio. The number of times arrangements of four blue beads were chosen is also greater than in the case of the 1 : 1 ratio but is very much below what we should expect on theoretical grounds. Finally, the number of sets with four yellow beads is considerably decreased, and less than the theoretically expected number. We see that the children here as well regard each set of four beads as a miniature of the population of beads in the urn.

The above Types I (a), I (b) and I (c) choices are presented in more detail side by side for comparison in Table 16.

Type II. Under Type II conditions there are 16 beakers (instead of four as in Type I), and as four beads at a time are drawn from the urn they are heard to drop into a beaker. The subjects record their answers to the question: how many blue beads are in the beaker? For some subjects the question is phrased: how many blue *and* yellow beads are in the beaker?

Table 16. Preferences for arrangements of beads under Type I conditions

Proportions of times out of 16, subjects choose the possible arrangements of four beads, each of which may be blue or yellow, when proportions of the two colours are as stated in column headings.

Arrangements	Proportions of blue and yellow beads in the urn stated to be				
	unknown	ratio 1 / 1		ratio 3 / 1	
		observed	expected	observed	expected
B B B B	1·59	0·67	1	1·74	5·0625
B B B Y B B Y B B Y B B Y B B B	3·95	3·67	4	8·49	6·75
B B Y Y B Y B Y B Y Y B Y B B Y Y B Y B Y Y B B	5·86	8·17	6	4·26	3·375
B Y Y Y Y B Y Y Y Y B Y Y Y Y B	3·36	2·70	4	1·40	0·75
Y Y Y Y	1·24	0·80	1	0·11	0·625
Ratio of blue to yellow beads, as chosen by subjects	1·08	1·04		1·95	
Number in sample Age	95 10+	30 11+		35 12+	

(or reversing the order: how many yellow and blue?)[1] The reader will take note that here the child is making a guess about the composition of a set of four beads. He is not, as in Type I, guessing the colour of *single* beads in sets of four. Here, too, three conditions have been studied (a) when the proportions are said to be unknown; (b) when they are said to be 1 : 1, and (c) when they are said to be 3 : 1.

II (a). *Proportions of blue and yellow beads stated to be unknown.* In the replies they give to our questions the children say whether any given beaker contains 0, 1, 2, 3 or 4 blue beads. If they wrote their answers down at random, their replies would yield an equal number of beakers containing 0, 1, 2, 3 and 4 blue beads. But they do not behave in this fashion.[2] Their choices imply that they regard two blue and two yellow beads in a beaker as more likely to occur than one blue and three yellow (or *vice versa*), and also more likely than no blue and four yellow beads (or *vice versa*). And they regard one blue and three yellow (*or vice versa*) as more likely than no blue and four yellow (or *vice versa*).

The statistically-minded reader will observe a characteristic difference between the results of Types I and II experiments respectively. In Type I, if the replies were random, the resulting distribution of arrangements would not be rectangular (i.e. an equal number of choices of arrangements of four beads with 0, 1, 2,

[1] In Experiment II (a) the question was: how many blue beads are in the beaker? In Experiment II(b) one group answered the question: how many blue *and* yellow beads?, and the other group, how many yellow *and* blue beads are in the beaker? The order II(a), II(b) was followed for one group and II(b), II(a) for the other.

[2] This can be shown to be the case by applying a test of statistical significance. The children's distribution of 0, 1, 2, 3, 4 blue beads differs significantly from a rectangular one and, indeed, shows a marked tendency towards a binomial distribution.

3 or 4 blue) but in the direction of a binomial distribution. In Type II, however, random choice would lead away from a binomial towards a rectangular distribution.

One feature of this Type II experiment should be specially noted. Slight variations in the total pattern of choices were obtained by the different forms of the experiment II (a) and II (b) respectively. These differences may possibly be due to the way the question was put. It seems that when we simply ask: "how many blue beads are in each beaker?" and do not refer to yellow beads, the children tend to avoid writing "no blue beads".

We conclude that, as in Type I, when the proportions of blue and yellow beads are stated to be unknown, children tend to allot blue and yellow beads in nearly equal proportions. Furthermore, they do not allocate proportions of blue and yellow beads *within sets of four* at random. They seem to have a significant liking for sets of two blue and two yellow beads, and they prefer an arrangement of one blue, three yellow (or *vice versa*) to no blue, four yellow (or *vice versa*).

Types II (b) and II (c). *Proportions of blue and yellow beads in the urn stated to be in the ratio of* 1 : 1 *or* 3 : 1. The chief effect of giving some information about the actual proportions of blue and yellow beads in the urn is to induce a distinct preference for saying that the proportion present in the urn occurs also in each beaker. The extent of the preference is shown by the frequency with which the proportions stated to be in the urn occur in the replies given by the sample of children as a whole.

The figures in the Table which follows relate to the averages of the entire groups. In fact, some 63 per cent. in each of the two groups (under 1 : 1 and 3 : 1 conditions) chose the proportions stated to be in the urn more frequently than any other proportion. This may perhaps be due to their treating each beaker as a miniature of the population in the urn, as in Type I (b) and (c). Nevertheless, other proportions of blue and yellow are by no means entirely absent. Moreover, the distributions on the whole have some slight resemblance to the distributions which would be obtained on purely theoretical grounds. This is borne out by the frequency with which certain features of the theoretical distribution occur in the distribution based on the values given by the children. When the proportions of blue and yellow beads are either assumed or known to be equal then, theoretically, the number of beakers with no blue and four yellow beads is the same as the number with no yellow and four blue. When, however, the proportions are either assumed or known to be in the ratio of 3 : 1 different frequencies of these proportions (no blue, four yellow and no yellow, four blue) are theoretically expected to occur. In fact we find that when blue and yellow beads are stated to be in equal proportions, significantly more children make the frequency of no blue, four yellow, the same as four blue, no yellow, than when the proportions are stated to be 3 : 1. And under the same condition (1 : 1), significantly more children tend to equate the frequency of the proportion one blue, three yellow, and three blue, one yellow, than when the ratio of blue to yellow beads is stated to be 3 : 1.

Finally, when blue and yellow are stated to be in equal proportions, the beads allotted by all subjects combined are in the ratio of almost exactly 50 : 50 as compared with a ratio of 62 : 38 when the proportion of blue to yellow beads is stated to be 3 : 1.

Table 17. Type II. Distributions of beakers derived from guesses when the information given to the children about the contents of the urn is varied.

Mean values as proportions of 16.

Information given to children	Distribution	Number of blue beads					Total
		0	1	2	3	4	
Proportions unknown $N = 186$ (age/ 10+)	Observed	1·8	3·5	4·3	3·6	2·8	16
	Calculated[1]	1·0	3·5	5·5	4·4	1·6	16
Proportions 1 / 1 $N = 35$ (age/ 12+)	Observed	1·1	3·6	6·4	3·8	1·1	16
	Theoretical	1	4	6	4	1	16
Proportions 3 / 1 $N = 30$ (age/ 11+)	Observed	0·6	2·3	4·4	6·4	2·3	16
	Theoretical	0·625	0·75	3·375	6·75	5·0625	16

[1] The calculated values have been obtained from the actual proportions of blue and yellow beads implicitly allotted by each child, not from hypothetical values of 50 per cent. blue and 50 per cent. yellow. The observed values are based on two experiments which varied in the form the question was asked.

Type III: Let us now examine the third way in which a frequency distribution may be built up. We begin by placing four beads drawn at random from the urn into each one of the 16 beakers. The subjects now guess the number of beakers containing 0, 1, 2, 3 or 4 blue (or yellow) beads, and these guesses *are* the distributions. The distributions are not derived indirectly as in Types I and II.

In previous studies we have found that the order in which a series of questions is asked in this type of experiment may affect the subjects' estimates, items in the earlier questions tending to be given higher values than they would otherwise get. This is itself a curious effect which deserves closer scrutiny. It would seem to mean that when a child has a given quantity to dispose of in several directions, he does not tend to dispose equal amounts in all directions. He tends rather to dispose of the largest proportion in the first direction and steadily diminishing proportions in the other directions. The estimates may also be affected by the order in which the colours are referred to in the questions. In order to eliminate these influences it was necessary to devise ten ways in which the questions were to be presented. In five of these, blue was mentioned first, and in five yellow was mentioned first. In the set of five forms of questions for each colour the questions were so arranged that all five forms appear in each position an equal number of times. The ten orders of presentation were given to approximately equal numbers of subjects, each subject receiving a single order of presentation. The means for the entire samples were obtained from the sub-group means.

A specimen set of questions is set out below :

NAME :................................ Age :..........

How many beakers will have four yellow beads and no blue beads?	....
How many beakers will have no yellow beads and four blue beads?	....
How many beakers will have three yellow beads and one blue bead?	....
How many beakers will have one yellow bead and three blue beads?	....
How many beakers will have two yellow beads and two blue beads?	....
TOTAL	16

Type III (a). *Proportions of blue and yellow beads stated to be unknown.* Our first task is to discover whether the children, when the proportions in the urn are stated to be unknown, estimate the number of beakers containing 0, 1, 2, 3 or 4 blue beads respectively in a manner that is actually likely to occur on *any* assumption that might be made about the contents of the urn. If, for example, a child stated that each one of the 16 beakers contained four blue beads, that would agree with an assumption that all the beads in the urn were blue. But if a child stated that half of the beakers did not contain a single blue bead and half contained four blue beads each, this would be most unlikely on any conceivable assumption about the contents of the urn. Oddly enough, one child gave just these values. We do not know whether the children make any explicit assumptions about the contents of the urn, but we may perhaps infer, from the number of blue and yellow beads indirectly assigned by each child that they do tend to assume roughly equal proportions. In spite of this, few of the individual distributions have a high mathematical probability of occurring in terms of the number of blue and yellow beads 'implicitly' assigned.

Bearing in mind that the subjects 'implicitly' allot approximately equal proportions to blue and yellow, it is of interest to see whether they tend to assign higher values to the number of beakers containing two blue and two yellow beads than to beakers containing any other combination of four beads. If they do, we could infer that they think of the samples in the beakers as miniatures of the contents of the urn. We find, indeed, that one subject stated that there were two blue and two yellow beads in every one of the 16 beakers. We can calculate the number of beakers allotted to two blue and two yellow beads and compare it with the number of beakers allotted to other combinations. We note that all three age groups assign the highest number to beakers containing two blue and two yellow beads. Details are shown in Table 18 which follows:

Table 18. Number of times subjects assign highest value to the categories of beaker with different numbers of blue (and yellow) beads.

Age	Number of blue beads in beaker				
	0	1	2	3	4
10+	8	10	18	9	7
12+	3	1	7	5	5
14+	2	1	5	2	4
Total	13	12	30	16	16

Let us now ask whether children at the ages 10+, 12+ and 14+ have any preference for assigning equal values to beakers with no blue and beakers with four blue beads, and/or equal values to beakers with one blue and three blue beads. To answer this question we can compare values assigned to these categories of beaker with those assigned to the remaining categories of beaker, taking them in pairs. This comparison is shown in Table 19.

The proportions in the above Table are based on the actual frequencies with which the various categories of beaker are chosen. The proportions for 0,4 and 1,3 are underlined in the columns for the 12+ and 14+ age groups. If we compare the frequency with which the number of beakers containing no blue beads is made the same as the number containing four blue beads, (and similarly for those containing one blue bead and

Table 19. Categories of beakers containing 0, 1, 2, 3 or 4 blue beads compared in pairs of category

$N = 87$ (10+), + 31 (12+), + 26 (14+)

Paired categories of beaker containing number of blue beads		Proportion in sample assigning *equal* values to pairs of categories of beaker[1]			Total
		10+	12+	14+	
0	1	0·17	0·16	0·12	0·16
0	2	0·17	0·13	0·08	0·15
0	3	0·22	0·10	0·19	0·19
0	4	0·16	0·42	0·42	0·26
1	2	0·15	0·23	0·23	0·18
1	3	0·17	0·45	0·35	0·26
1	4	0·26	0·10	0·12	0·20
2	3	0·15	0·16	0·23	0·17
2	4	0·25	0·13	0·12	0·20
3	4	0·15	0·13	0·12	0·14

[1] The proportion of unequal values assigned to each pair of categories may be obtained by subtracting the values given from 1·00.

three blue beads) with the frequency with which other pairs of category (e.g., one blue bead and four blue beads), we find a marked difference. The former two pairs (0,4 and 1,3) are chosen equal by the children aged 12+ and 14+ significantly more often than other pairs of category. The ten-year-olds do not show this tendency.

The twelve-year-old children in one of the experiments

($N = 30$) were asked to give reasons for making their particular choices. The features just described are brought out clearly in the reports given by the children, as quoted below:

Subject (1)

The beakers with one yellow and three blue, and those with three yellow and one blue, have opposite amounts, so I've made them the same. And I've done the same as that with the ones with four blue and no yellow and four yellow and no blue.

	No. of blue beads				
	0	1	2	3	4
Distribution actually given	4	2	4	2	4

Subject (2)

There will be most in number V (2B, 2Y) as there is an even chance for each two. Numbers I (4B, 0Y) and II (0B, 4Y) rule each other out and so I have put them even. Numbers III (3B 1Y) and IV (1B 3Y) will rule each other out also. As there is not an even number left I simply put four for number III (3B 1Y) and three for number IV (1B 3Y)[1]

	No. of blue beads				
	0	1	2	3	4
Distribution actually given	2	3	5	4	2

Subject (3)

I thought that the beakers having an equal number of beads in them would not be many and I only put one. Also the first and second answers (0B, 4 Y and 4B, 0Y) are likely to be about the same, and the third and fourth answers (1B, 3Y and 3B, 1Y).

	No. of blue beads				
	0	1	2	3	4
Distribution actually given	4	3	1	2	6

[1] This subject herself numbered the questions I to V.

The children, as we have seen, believe that beakers with two blue and two yellow beads occur most frequently among the 16 beakers. Let us now see whether they also believe that beakers with one blue bead are more frequent than beakers with no blue beads and beakers with three blue beads more frequent than those with four blue beads. They would presumably do so if they had some idea of the relative probabilities of the different possible combinations. We can compare the frequencies with which these beliefs are represented in the numbers assigned to the different categories of beaker. In fact, we find that all the three age groups (10+, 12+ and 14+) tend to assign greater values to beakers with no blue beads than to those with one blue bead; and similarly greater values to beakers with four than to beakers with three blue beads. We might account for this behaviour in terms of the marked preference for homogeneous and the avoidance of asymmetrical arrangements of beads as shown above in the Type I experiments. It is, however, a tendency away from, rather than towards, a binomial distribution.

Type III (b) and III (c). *Proportions of blue and yellow beads the urn stated to be in the ratio of* 1 : 1 *or* 3 : 1. Let us take the three age groups in turn. In the two younger groups, whose ages extend from 10 to 13 years, there is little if any change in the distributions as a result of providing information about the contents of the urn. But at the age 14+ we can begin to detect some effect. When the proportions are stated to be 1 : 1 the average values assigned to the beakers with different numbers of blue beads tend much more clearly towards a theoretical distribution than when the proportions of blue and yellow are stated to be unknown. And when the two colours are stated to be in the ratio of 3 : 1, the number of beakers believed to contain 0, 1, 2, 3 or 4 blue beads respectively are in the same order of magnitude as the theoretical numbers, although the numbers themselves are attenuated. The mean value allotted to beakers with three blue, one yellow, under these conditions is 4·5, which is the highest value obtained for this proportion of colours in all the Type III experiments.

In Tables 20, 21 and 22, the reader will find the distributions given by the children under conditions Type III (a), (b) and (c), i.e. when proportions of blue and yellow beads in the urn are stated to be unknown, in the ratio 1 : 1, and in the ratio 3 : 1. The values assigned by each age group are shown in a separate Table. In each Table, the numbers allotted by the children to the categories of beaker are expressed as proportions of the total number of beakers (16).

Table 20. Type III. Children's distributions of beakers containing 0, 1, 2, 3 *or* 4 *blue beads when the information about the contents of the urn is varied. Mean values allotted as proportions of* 16.

Age 10+[1] $N = 86$

Information about proportions in urn	Number of blue beads					Total number of beakers
	0	1	2	3	4	
Unknown	3·2	3·1	3·7	2·9	3·2	16
1 / 1	3·2	3·2	3·7	3·0	3·0	16
3 / 1	3·2	3·0	3·4	3·4	3·0	16

[1] Including some eleven-year-olds.

Table 21. Type III. Children's distributions of beakers containing 0, 1, 2 3 or 4 blue beads when the information about the contents of the urn is varied. Mean values allotted as proportions of 16.

Age 12+ $N = 31$

Information about proportions in urn	Number of blue beads					Total number of beakers
	0	1	2	3	4	
Unknown	3·3	2·7	3·5	3·2	3·1	16
1 / 1	3·0	3·7	3·7	3·7	1·9	16
3 / 1	2·8	3·3	3·1	3·0	3·6	16

Table 22. Type III. Children's distributions of beakers containing 0, 1, 2, 3 or 4 blue beads when the information about the contents of the urn is varied. Mean values allotted as proportions of 16.

Age 14+ $N = 26–35$

Information about proportions in urn	Number of blue beads					Total number of beakers
	0	1	2	3	4	
Unknown	3·0	3·0	3·7	3·2	3·2	16
1 / 1	3·0	3·7	4·0	3·1	2·7	16
3 / 1	1·8	2·6	3·4	4·5	3·7	16

The time has come now to review these three types of experiment. First of all a word of justification is needed for the phrase 'implied assumptions' which we have occasionally used. This phrase is not intended to mean that children necessarily make any assumptions. It refers simply to the total number of blue and yellow beads actually allotted by each child. Any such number presupposes that some idea about the contents of the urn has been taken for granted, for example, that the urn contains *some* blue and *some* yellow beads.

In all three types of experiments, when the proportions of blue and yellow beads in the urn are stated to be unknown, the totals allotted do not diverge much from roughly equal numbers of blue and yellow beads. The children could, of course, have given values based on any one of a very large number of possible ratios of blue and yellow. When actually asked to say what they thought were the proportions in the urn, about 10 per cent. said that it contained equal numbers of blue and yellow beads.

When the proportions are unspecified there seems to be a tendency for children, and perhaps adults too, to assume that the proportions are equal or nearly so. The assumption of equality, under the conditions described, seems to be the simplest and the most naturally acceptable.

These 'implicit' estimates, when the ratios in the urn are not specified, resemble corresponding estimates made in these as well as other experiments by comparable groups of children when the proportions in the urn are stated to be 1 : 1. When, however, the proportions are specified as 3 : 1 the 'implicit' estimates are quite different.

Again, in the three types of experiment, when the proportions are stated to be unknown, the children's choices, at least the majority of them, are not of the kind that could have arisen in random fashion. In Type I there is a clear preference for certain kinds of arrangement. In Type II the children do not tend to allot similar proportions of blue and yellow beads to the different beakers, a result which would be expected to occur on a random basis. They expect more beakers to contain two blue and two yellow beads than no blue, four yellow. In Type III the choice of numbers which add up to 16 is likewise not random, and here too there is a preference for saying that there are many beakers with two blue and two yellow beads.

So far we have been discussing the average values given by the groups of subjects as a whole. We can perhaps better understand the states of mind of individual children if we study the explanations given by a group of 30 twelve-year-olds. We do not, of course, suppose that their explanations fully account for the choices actually made by them. Some 20 per cent. of the group appear to have some idea of a statistical distribution, and they mention one or more reasons for the various effects which have emerged in the analysis. Some 40 per cent. offer some definite reason for the way in which they have made their choices. They say, for instance: "I wanted to have every number different" or "I did not think any of the numbers would be alike." The remaining 40 per cent. gave no definite reason for their choices. One said: "They were the first numbers which came into my head and added up to 16. When I had done the first three I divided what was left into two as I thought best."

Briefly the effects of varying the information and of specifying the proportions in the urn as 1 : 1 or as 3 : 1 are as follows:

(a) Type I. When we tell the children that the proportions in the urn are 1 : 1, they increase the number of arrangements with two blue, two yellow beads. When we tell them that the proportions are 3 : 1, they increase the number of arrangements with three blue, one yellow.

(b) Type II. The same effect occurs as in Type I, but here the effect appears in relation to the contents of each of the 16 beakers and not, as in Type I, in relation to arrangements of sets of four beads. A significantly larger number of children say that the number of beakers with no blue, four yellow is the same as that for no yellow, four blue under the 1 : 1 condition than under the 3 : 1 condition. Exactly the same effect occurs with one blue, three yellow and *vice versa*.

(c) Type III. One effect of stating the ratio to be 3 : 1 is to change the total numbers of blue and yellow beads allotted as compared with the condition 1 : 1. There is very little difference between the total number of blue or yellow beads allotted under conditions 'unknown' and 1 : 1, but under condition 3 : 1, the fourteen-year-olds definitely allot a larger total of blue beads than yellow beads. The superiority of the fourteen-year-olds in this respect is brought out in Table 23.

Table 23. Ratio of blue to yellow beads in the totals allotted.

Age	Proportions in urn stated to be		
	Unknown	1 / 1	3 / 1
10+	1 : 1·00	1 : 0·96	1 : 1·00
12+	1 : 1·00	1 : 0·88	1 : 1·06
14+	1 : 1·06	1 : 1·00	1 : 1·43

A second effect appears in the distributions given by different age groups. Only at the age 14+, in our groups, does the information 3 : 1 begin to have an effect on the character of the distributions themselves.

Finally, let us look at our conclusions from the point of view of the general intellectual development of the children. There appears to emerge, between 10 to 15 years, four stages in the understanding of the idea of a distribution. The first glimmering of the idea is shown in a belief that the numbers in the set or distribution will vary. This is followed later by a tendency to assume that the category of equal proportions (two blue, two yellow) will occur most often, although a few children seem to think that it will occur least often. At the third stage the same likelihood is assigned to categories with similar structure (number of beakers with one blue, three yellow = number with three blue and one yellow; and similarly with no blue, four yellow and *vice versa*. Even at the age of ten, children tend to believe that beakers with two blue, two yellow are more likely to occur than other combinations. But it is at the age of twelve years that equal values are first assigned to beakers with no blue, four yellow, and no yellow, four blue; and similarly for one blue, three yellow and *vice versa*. This last preference may be an important new step beyond the level of the ten-year-olds for it makes a *symmetrical* distribution possible. The tendency to symmetry may also illustrate the use of a 'principle of indifference'. A further stage, which is apparently not yet reached at the age of fifteen in our groups is to assign, when the ratios in the urn are not known, greater likelihood of occurrence to one blue, three yellow than to no blue, four yellow; and similarly greater likelihood to three blue, no yellow, than to four blue, no yellow.

We venture the hypothesis that the stages we have distinguished in the development of the idea of a distribution partly represent mental 'structures' (perhaps in Piaget's sense of the word) that emerge at different ages and partly represent the fruits of experience. These 'structures' undergo changes during development. Shaped at the start by purely subjective preferences of one kind or another, they become less and less controlled by such subjective influences and assume a form which is more in accord with the objective situation.

Chapter Eight

SUBJECTIVE PROBABILITY IN GAMBLING

The impetus given by gambling to the study of *mathematical* probability makes the behaviour of the gambler particularly relevant to the study of *subjective* probability. When we gamble we exemplify *par excellence* those tendencies which find in subjective probability and risk-taking their most characteristic expression. The various experiments described above may shed some light on the mind of the gambler. Of special interest is the way we judge the outcomes of events as independent.

The practice of gambling has always been extraordinarily widespread. It is not a local or passing craze which can be snuffed out by a few vigorous sermons. The loftiest specimen of such a sermon is the one entitled *The Consequences of the Vice of Gaming as they affect the Welfare of Individuals and the Stability of Civil Government, Considered*, preached by Prebendary Thomas Rennell at Winchester in 1794. The Prebendary speaks of gaming as a subject of "the highest, most awful, and most instant importance".

Gambling is one of the most characteristic features of human activity everywhere and common to the most diverse cultures. In the widest sense, it is practised by people of all ages and enters, in one form or another, into their work and play alike. A contemporary writer claims that gambling "... is undoubtedly a spiritual affair . . . no man would gamble unless he had a love of life and relished what he could obtain from life with his winnings".[1]

Gambling is found in the most far-away times and places among highly civilized and primitive peoples alike, in the ancient civilizations of China (once described as "one vast gaming-house"), Japan, India, Persia, Egypt, Palestine, Greece and Rome, and among peoples as scattered as the American Indians, some of whom believed that gambling would be continued to be enjoyed in the after life, the Zulus, the Malays and Javanese. In some legends of the Hindus, Rajahs play for days on end until the loser is reduced to slavery or has to go into exile.[2] The Emperors Augustus, Caligula, Claudius, Nero and Domitian were counted among the most passionate of gamblers. Nearer home we find in Germany from the earliest times and in France many people in the grip of gambling. In the France of Henry IV hardly anyone escaped the frenzy of gambling in spite of strict legislation against it. When Louis XIV died, most French people are said to have thought of nothing else but gambling. In England, dice-playing dates from the Roman occupation. A twelfth-century edict shows that gambling was very common at that time, and Ordericus Vitalis records that bishops and

[1] Mills, W. G., 'The Spirit in Gambling', *Sporting News*, Issue No. 3, (Second Season), London: 24 September 1953.

[2] France, C. J., 'The Gambling Impulse', *Amer. J. Psychol.*, 1902, 13, 364–407. He quotes Juvenal, a contemporary of Domitian, who writes: "When was the madness of games of chance more furious? Nowadays not content with carrying his purse to the gaming table, the gamester conveys his iron chest to the playroom. It is there you witness the most terrible contests. Is it not madness to lose one hundred thousand sestertii and refuse a garment to a slave perishing with cold?"

other clergymen were fond of dice-playing. Ordinary people were prohibited by Henry VIII from gambling except at Christmas. We learn something about gambling in the England of the seventeenth and eighteenth centuries from Theophilus Lucas' *Memoirs of the Lives, Intrigues and Comical Adventures of the most Famous Gamesters and Celebrated Sharpers in the Reigns of Charles II, James II, William III and Queen Anne* (London, 1714). A correspondent in the *Grub Street Journal* complains that "the canker of gambling is surely eating into the very heart of the nation". Towards the end of the eighteenth century many old-established gambling houses in London, which had been closed, were reopened. These were kept by women, to whom gambling appealed as much as to men. During the nineteenth century, writes Ashton,[1] "the West End of London literally swarmed with gambling houses". It was also a period of uncurbed speculation exemplified in the South Sea Bubble and the dozens of other wildly extravagant enterprises.

The stakes were often exceedingly high. In China, we are told, they sometimes staked their fingers or the hairs on their head. St. Ambrose tells us that among the Scythians it was common to risk one's life on a single throw of the dice. Some have staked their teeth, eyebrows, their wives, their families or their freedom. One gambler stated in his will that his skin and membranes should be used to cover a table, dice box and draught board, and that his bones should be converted into dice.

Gamblers often get worked up into a frenzy. They eat the cards, crush the dice and damage the furniture. One infuriated

[1] Ashton, J., *History of Gambling in England.* London: Duckworth, 1898.

player jammed his mouth with a billiard ball and it stayed there until removed by a doctor. A Neapolitan player is said to have bitten into the table with such force that his teeth entered the wood and there he remained until he died.

The strength of the gambling impulse in Britain to-day may be roughly gauged by the following facts.[1] About 16 million people gamble in football pools and about one in four of these does so regularly. Approximately half a million people attend dog-racing meetings weekly. About 6½ million people bet on horses. This leaves out entirely all the gambling carried out privately, illegally, in small communities, and in pin-table saloons, church bazaars, and fun-fairs. It also leaves out the world of speculative investment.

Capital, Gain and Investment

The study of gambling need not be limited to laboratories. Certainly parents do not encourage their children to take part in experimental gambling. They take this view in spite of the fact that such preparation might make their offspring more realistic at the mellow age of twenty-one when they can forthwith commence to fill in their own football pool coupons. Fortunately for us we have in this country a public Corporation blessed with breadth of vision and a traditional sporting spirit, which conducts gambling sessions and encourages the public to participate in them.

Several pertinent questions arise in monetary gambling and B.B.C. broadcasts provide suitable material for examining, at least in a preliminary way, some of these questions. What

[1] Based mainly on the *Report of the Royal Commission on Gambling*, H.M.S.O., 1951.

relationship is there between the amount of capital one possesses and the amount one is prepared to stake? How does a win or loss in a given stake affect subsequent stakes?

In two broadcasts to the Forces, on the 10th and 23rd of June 1953, members of the studio audience were invited to take part in a "quiz". They were given an initial capital of ten shillings and told that they would be asked three questions. They could invest part or all of their capital on the first question. If they answered correctly they retained their stake and won an equal amount. If they answered incorrectly, they forfeited their stake. They could then invest in a second question under similar conditions, and so on with the third question. The maximum which a player could take away at the end of the session was £4, including the initial ten shillings.

Before each question, the quiz-master usually indicated how difficult the question was. Nine members of the studio audience participated during the two sessions. The behaviour of these nine players is summarised in Table 24.

The degree of difficulty of the questions is represented in the Table by the letters A to E. We have used these letters to convey our impressions of the quiz-master's remarks. Where no letter is given, no remark was made.

A—very easy
B—easy
C—neutral
D—difficult
E—very difficult.

In answer to the first question raised above, there certainly seems to be a relationship in these circumstances between the

Table 24. Gambling in B.B.C. Broadcasts to the Forces

Number of Participant	Number of Question	Information given by Quiz-master	Capital in hand (shillings)	Amount invested (shillings)	Result: win or lose
1 (Wren)	1	B	10	4	win
	2	B	14	6	lose
	3	—	8	4	lose
2 (Soldier)	1	B	10	6	win
	2	D	16	10	win
	3	—	26	13	win
3 (Airman)	1	D	10	6	lose
	2	C	4	2	win
	3	D	6	4	win
4 (Wren)	1	D	10	3	win
	2	D	13	9	win
	3	D	22	15	win
5 (Male Civilian)	1	C	10	4	win
	2	B	14	6	lose
	3	B	8	6	win
6 (Wren)	1	C	10	2½	win
	2	B	12½	9½	lose
	3	—	3	3	win
7 (Airman)	1	B	10	5	win
	2	B	15	8	win
	3	E	23	15	lose
8 (Wren)	1	—	10	5	win
	2	B	15	7	lose
	3	B	8	4	win
9 (Male Civilian)	1	—	10	6	win
	2	D	16	6	win
	3	D	22	6	lose

amount of capital one possesses and the amount one stakes. The relationship is such that the stake increases with size of capital, and the proportion of stake to capital remains fairly constant at about 60 per cent.

The players were familiar with the kind of questions asked and would therefore have some idea of their chances of winning. We should not expect a proportion of 60 per cent. in other situations where the difficulty of the task might be much greater. This proportion may be thought of as an index of the player's subjective probability of success.

We may answer the second question by saying that the effect of a win is to increase the stake, and of a loss, to decrease it. On all the 13 occasions on which a player won, he never subsequently lowered his stake, and in 11 of these he raised it. On the five occasions when a player lost, in no instance did he increase his stake and in four of these he decreased it. In certain cases the player is forced to lower his stake because of reduced capital following a loss.

We are not, of course, suggesting that any firm conclusions can be based on such slight evidence. The details are given to suggest methods whereby gambling tendencies may be more fully explored. We express the hope that the B.B.C. will continue to sponsor such sessions and allocate part of its budget to the advancement of our understanding of gambling.

Winning Football Pools

Gambling practices vary from situations, at one extreme, where no skill whatsoever can effect the outcome, to situations at the other extreme, where skill plays the decisive part. An example of the first is a well-conducted lottery; an example of the second is a golf tournament. Intermediate between these two extremes there is a wide range of practices into which skill and chance enter in different combinations. Success in many card games depends on the skill of the player and the cards he happens to receive. Where do football pools stand on this scale? We might expect that the result of a football match depends on the skill of the teams and that this skill will be reflected in the record of their past performances. We still have to ask whether the most expert of football enthusiasts can make successful predictions. Newspapers employ experts to predict the results of matches and their predictions usually agree with one another to a fair extent. But few people ever try to check these predictions or compare the accuracy of the different experts. It is known that about half of all matches result in a victory for the home team, a quarter end in a victory for the 'away' team, and a quarter are drawn. If the reader takes all matches played in a given week and predicts their outcome in a random manner, allocating them in these proportions, he could check the accuracy of his predictions and compare them with those of the experts. He will find that the experts exhibit little superiority over his own chance guesses. Football pool investors widely believe that the pool promoters are skilful in selecting likely drawn matches and that they include these in a particular pool where the investor has to predict the results of a limited number of matches[1] (12 or 14)

In an analysis[2] of matches in a period of 51 weeks it was

[1] Penny Points Pool.

[2] Carried out by Mr A. Stuart and reported in the *Journ. Roy. Stat. Soc.*, 1952, 115. II, p. 215. The analysis related to Littlewood's Penny Points Pool.

found that 20 per cent. of the matches in one such pool resulted in a draw whereas the proportion of draws in the matches not included in this pool was 22 per cent.

Although it has been amply demonstrated that there is little, if any, skill in predicting the outcomes of football matches, this need not deter the eager student of subjective probability from engaging in such pursuits. It is true that he is unlikely to be more successful than his fellows in predicting the results of matches, but that is not his aim. His single aim and object is to receive from the pools more than he invests in them. We suggest that the most promising way to achieve this goal is to turn his skill in a different direction.

Consider the following facts: (a) the results of the matches are unpredictable; (b) the pool investor is playing against his fellow investors in that the total sum invested is shared among them all; and (c) some of the population hold a firm belief that the results of certain matches are predictable on the basis of past results and sometimes they follow the experts' advice. Now let us take a simple pool where the investor has to select three drawn matches from about 60 which are to be played. As all matches on the average have an equal chance of being drawn, any choice of three matches as drawn is as good as any other. The prize-winning potentiality of different choices may, however, vary considerably because it depends on the number of investors making that particular choice. So a choice of three matches strongly recommended by the experts has an equal chance of being accurate as any other choice, including a choice of three matches which all experts agree will not be drawn. Accordingly, the greatest return from an investment may be expected if the investor chooses matches as draws which have the least likelihood of being chosen by other investors as draws. The student of subjective probability should therefore seek those pools in which the experts themselves invest and then carefully avoid following the expert's advice. If he perseveres for an adequate period he may have the satisfaction of finding that his studies are rewarded. This conclusion holds good, however, only if studies of subjective probability are not too widely disseminated.[1]

Gambling in Relation to Destiny, Divination and Primitive Justice

The distinctive characteristic of gambling, its concern with unpredictability, has caused it to be closely connected with ideas of destiny and divination and with the development of the concepts of primitive justice.

The entire world, in the Mahābhārata, is treated as a game of dice in which Siva plays with his queen.[2] The Ases of Eddic mythology determine the fate of the world by casting dice. In the *Iliad*, Poseidon, Zeus and Hades share the world between them by shaking lots, and these lots are credited with a virtue or consciousness of their own whereby they ensured a correct decision, or believed to be controlled by higher

[1] We can calculate that if in a particular pool the proportion of the total investment returned as prize money is W and the proportion of the public following the experts is R, the mean ratio of winnings to investments for these investors will be W/R, and for those who do not follow the experts the mean ratio will be $W/(1-R)$. Thus, if $W = 0 \cdot 6$ and $R = 0 \cdot 9$, the ratio for those following the experts will be $0 \cdot 67$ and for the others $6 \cdot 0$.

[2] Held, G. J. The Mahābhārata (Thesis), Leyden, 1935 (quoted by Huizinga, *op. cit.*).

powers whose will they revealed. Zeus is expected to control the lots which decide which of the Greek heroes is to fight with Hector.[1]

There are countless instances of gambling as divination among ancient and primitive peoples. The Romans turned their *astrali* into dice by putting numbers on the sides. The Maoris threw lots to identify a criminal. In Samoa a nut was spun round for the same purpose. In the Tonga Islands nut-spinning was at once an amusement and an art of divining whether a sick person would recover. A Zande technique[2] is to administer a specially prepared substance to a chicken while asking a question. The chicken vomits the substance or is killed by it, and so answers "yes" or "no". We, ourselves, use the same cards for playing and for telling fortunes.

But divination by lot, dice, nut-spinning or other method among primitive peoples is never a resort to mere unpredictable chance. It is an appeal to a supernatural agency to decide the issue. More sophisticated peoples, too, believe that the gods turn the dice or shuffle the cards. Indeed, Tylor[3] many years ago expressed the view that the gambling of primitive man may have originated as a method of divination, the same implements being used for both. Our own games of chance may be survivals from such practices. Some of our economic judgments may retain something of this divinatory character. We are told that in Throgmorton Street, the heart of the

[1] See Onians, R. B., *The Origins of European Thought*, pp. 392–3. London: Cambridge University Press, 2nd edition, 1954.

[2] Evans-Pritchard, E. E., *Witchcraft, Oracles and Magic among the Azande of the Anglo-Egyptian Sudan*, Oxford: Clarendon Press, 1937.

[3] Tylor, E. B., *Primitive Culture*, I, pp. 78–83. London: Murray, 1871; see also Culin, S., 'Chess and Playing Cards', *Report of the Smithsonian Inst.*, 1896.

stockbroking world in London, a man used to sit with a bag of nuts. Passers-by drew a handful from the bag. If they guessed the number of nuts correctly they would receive a penny for each nut; if not, they paid a penny. Many stockbrokers regulated their 'bulling and bearing' by this method of divination.[1]

Ideas of primitive justice are associated with the principle of retribution (*lex talionis*) and the contest. Retribution embodies the notion of balance or symmetry: a disadvantage sustained by one must be repaid with the same disadvantage, and similarly with an advantage. This is exemplified in 'riddle myths' where the one who cannot solve the enigma is punished by death or in some other way, and he who solves it gets a prize. Yājñavalkya, the riddle-solving priest of the Brāhamanas, is so effective that he makes his opponent's head drop off. These are forms of retribution; just as vice deserves punishment so virtue deserves reward. In the so-called 'wager tales' as well, especially those about winning a race, we find the same idea. There are many forms of the story of the race between the hare and the tortoise. The hare mocks at the tortoise who challenges him to a race and achieves victory as a reward for his persistence.[2] In the contest, however, the decisive element is determination by the gods of an outcome unpredictable by man. Huizinga[3] has suggested that the lawsuit was originally a contest, and the contest was the starting point for judgment by ordeal. Winning as such was the proof of truth and justice. Riddles, contests, ordeals, and verbal battles such as those of the Eskimos, were, like games of

[1] France, *loc. cit.*

[2] Kelsen, H., *Society and Nature*. London: Kegan Paul.

[3] Huizinga, J., *Homo Ludens*, p. 117, London: Kegan Paul.

chance, all ways of discovering sacred decisions of the gods.[1]

It has been suggested that this interpretation of human or animal situations was carried over into the realm of natural phenomena. The ideas of order, justice, and punishment were, it is supposed, borrowed by the Greeks from the domain of law and applied to the universal processes of nature. The order of nature was seen as part of the social order and natural processes interpreted in terms of a lawsuit.[2]

The Gambler

We may suspect that the typical gambler is moved by a powerful inclination to favour as the outcome of a future event, that which previously happened relatively rarely. When he places a bet on a horse he may keep on backing the favourite in each race in the belief that eventually a favourite must win. Another gambler may decide, after a succession of losses, that his luck is bound to turn. His hopes would be reinforced if he shared the view of many children that luck and ill-luck are stores which may be depleted or run out altogether. Each behaves as if he follows a principle of depen-

[1] Professor Jules Prussen has written to us privately that this view "may lead to the conception of the cosmic world as a play-board (powerfully expressed by Heraclitus and also by Plato, *Laws*, X, 903d). It seems that the two trends (i.e. retribution and unpredictability) primitively intermingled until there arose the conflict between the 'natural' notion of Cosmos and the 'unnatural' anxious feeling, the outcome of which was supposed unforeseeable, not on account of an objective independence of events, but on account of a higher ability of the mischievous partner in concealing its designs."

[2] Kelsen, H. *op. cit.*, The word 'etios' originally meant guilt before the law but later assumed the meaning of natural cause.

dence, the fall of the die being for him bound by the series of outcomes which has already occurred. But it seems to be a one-sided dependence. It leads him to favour change to the outcome that has happened less often in the past and it is not balanced by a belief in the continuance of the previous outcomes. His hopes of ultimate success may even rise with a mounting number of failures. Sometimes the gambler's subjective probability of eventual triumph may be so high that far from being a risk-taker, he hardly takes risks at all. In most gamblers the subjective probability of success is much higher than the mathematical probability.

The children in our experiment who have some idea of independence believe that guessing has nothing to do with intelligence. "It is not intelligence that counts in guessing" writes one. Those who lack this idea and are inclined instead towards ideas of independence seem to associate successful guessing with shrewdness. It may be that gamblers resemble this second group, for the gambler often thinks his otherwise inexplicable success is due to special sagacity on his part, as if he were in possession of occult powers in the form of an inner divining-rod. Luck he equates with cleverness, unluck with stupidity.

These feelings are vividly described by a man who played at Monte Carlo. He wrote afterwards:

> I was distinctly conscious of partially attributing to some defect or stupidity in my own mind, every venture on an issue that proved a failure; that I groped within me for something in me like an anticipation or warning (which, of course, was not to be found) of what the next event was to be, and generally hit upon some vague impulse in my own mind which determined me; that when I succeeded I raked up my gains, with a half impression that I had

been a clever fellow, and had made a judicious stake, just as if I had really moved a skilful move at chess; and that when I failed, I thought to myself, "Ah, I knew all the time I was going wrong in selecting that number, and yet I was fool enough to stick to it", which was, of course, a pure illusion, for all that I did know the chance was even, or much more than even against me. But this illusion followed me throughout. I had a sense of *deserving* success when I succeeded, or of having failed through my own wilfulness, or wrong-headed caprice, when I failed.[1]

He observed that one of his fellow-players looked on himself as a very sagacious person and received the respect of others as if he were in possession of occult powers. Once he heard that the youngest of his companions won at the last moment and retrieved his losses by putting his only coin, a 2-franc piece, on the number representing his age. His respect for the young man's shrewdness rose, and he reproached himself for not having done the same. He regretted not having put money on numbers with some relation to his life, the month of his birth or the number of the house he lived in.

We find then that the ideas of destiny, divination and the manner of dispensing justice appear to arise in connection with the contemplation of unpredictable events. The manner

[1] Letter to *Spectator*, 24 October 1873. Freud has interpreted the passion for gambling as a substitute for the compulsion to masturbate. He draws an analogy between the two practices. In both forms of 'play' there is an excited activity of the hands; in both, desire cannot be resisted; in both, all the promises never to do it again are always broken; the pleasure and the subsequent guilt and remorse are common to both. In gambling, he suggests, as in masturbation, it is the game for its own sake—*le jeu pour le jeu*. (Freud, S. Dostoevsky and Parricide, *Collected Papers*, London: Hogarth Press, 1950, Vol. V, pp. 222–242.

in which our subjects explain their reasons for choosing the outcome of an independent event suggests that in their minds too these ideas are interrelated. Some of them appear to believe that the outcome of a toss is determined but unpredictable, that it reveals the superiority of the successful guesser, and that a balanced number of different outcomes is fair and just.

Finally, the whole course of evolution may be looked upon, in a sense, as a gigantic gamble, a point of view implied in a recent study by Grey Walter. He suggests, first, that the brain must guess, during its selection of stimuli, whether a signal is, so to speak, worth bothering about. The brain reckons "the odds in favour of one event or one set of events implying another. In the simplest, isolated case, as in the laboratory, the odds are based on form—on the past history of that particular twin set of events being considered, on how they have come in before. But in more complex natural conditions the resemblance to the race-course becomes even closer, for a creature may have a wide choice of events and like any gambler can back a short odds favourite or take a longer chance on an unpopular outsider. On the favourite—that is, on the events that are obviously related—all backers have some money; the rewards are small but fairly certain. On the long chance there is less competition, more risks, a longer wait, more anxiety, but the hope of a fortune. In the evolution of living creatures can be seen the signs and effects of the struggle for the daily double. The animals that are well adapted and specialized—worm or fish, mighty lizard or mastodon, those who for some epochs were lords of the earth—were cautious backers of the dead cert, those whose nervous systems lay fallow during long ages, subdued by their mighty neighbours, starfish or newt, little furry creature or heavy-headed ape, were all the time quietly developing a system to pick out dark horses".[1]

In conclusion, we suggest that the foregoing studies may be regarded as offering an avenue of studying novel aspects of thought. They may also form a new link between studies of human and animal behaviour, for many of the experiments can be adapted for animal subjects. And they may have a more general implication. Now and then a psychologist is perturbed at the idea that his 'discipline' may not be looked upon as a science and is regarded rather as the shadow of a shadow of a science, and he is at pains to convince himself and the rest of the world that it really is a science. One way of self-assurance is to introduce the word 'scientific' in the titles of publications. But this is an expression of needless anxiety. The best evidence that the operations of mind obey certain laws and that a science of the mind is therefore possible lies in the fact that other sciences do exist. As Boole[2] said a hundred years ago, if the mind did not obey laws of its own it could never arrive at natural laws. One of our tasks should therefore be to seek to determine the laws of mind by tracing the path, stage by stage, along which we gradually move away from subjectively determined patterns of thought towards forms which are in closer accord with the objectivity of the order of nature and the logic of probability. The study of subjective probability is intended to be a small step in this direction.

[1] Walter, W. Grey, *The Living Brain*, p. 114. London: Duckworth, 1953.

[2] Boole, G., *The Laws of Thought*, p. 2. London: Walton and Maberly, 1854.

INDEX

DATE DUE SLIP