1-1-1983

Conceptions of probability: reality between a rock and a hard place.

Clifford E. Konold
University of Massachusetts Amherst

CONCEPTIONS OF PROBABILITY:
REALITY BETWEEN A ROCK AND A HARD PLACE

A Dissertation Presented

By

CLIFFORD E. KONOLD

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

February 1983

Psychology

This research was supported by NSF Grants

SED-8113323 and SED-8016567

CONCEPTIONS OF PROBABILITY: REALITY BETWEEN A ROCK AND A HARD PLACE

A Dissertation Presented

By

CLIFFORD E. KONOLD

Approved as to style and content by:

Alexander Pollatsek

Alexander Pollatsek, Chairperson of Committee

Arnold D. Well

Arnold D. Well, Member

Harry Schumer

Harry Schumer, Member

George Cobb

George Cobb, Member

Bonnie R. Strickland

Bonnie R. Strickland, Department Head
Psychology

ACKNOWLEDGEMENTS

I was riding the PVTA bus from Amherst to Northampton a few months ago, chatting with my favorite bus driver, Rob Taylor. He posed the following question -- "How many graduate students does it take to screw in a light bulb?" Not realizing how well I knew the answer, I asked how many. "Only one," he said. "But it takes seven years." I forced a laugh and thought about the past seven years.

Several individuals have played important roles for me during my graduate career, and I thank them--

My committee, Sandy, Arnie, Harry, and George, were as friendly a crew as their names suggest. I am especially grateful to Sandy. I've never worked with anyone so brilliant and felt so unthreatened. He both encouraged and then helped to clarify my thoughts. And he found inoffensive ways to motivate me. I'm also indebted to George, not only for the enthusiasm he always communicated to me about this study, but for his many insightful suggestions, most all of which I used.

Though neither of them worked directly with me on this project, I continue to be influenced by Mike Royer who remains my imaginary audience, reminding me to say it simply and clearly, and by Howard Gadlin who has showed me new dimensions of inquiry.

It would be impossible to express here the importance of the friendship of Doug Frost, Dale Thompson and Brian Stagner. Over the phone, Brian talked me down from the rafters more than once; and Doug and Dale provided the ballast that allowed Brian to get some needed sleep. These are the kind of friends that inevitably move away.

No one is more relieved with the completion of this dissertation than Amy Robinson. One morning, crippled by the fear (perhaps it was a hope) that I would die before I could finish the writing, I asked if she would complete it -- she had been a part of every phase of the work. We played catch with each idea and planned strategy together. She read through every draft and filled each with pointed criticism and ingenious rewordings. And she kept me energized with her love.

Finally I thank the individuals who allowed me to interview them. I encounter several of them with some frequency, but they seem to not remember (or want to forget) that they knew me. But having lived with their videotaped voices and faces, and spun theories around their utterances, they have become like old roommates.

ABSTRACT

Conceptions of Probability:

Reality Between a Rock and a Hard Place

(February, 1983)

Clifford E. Konold, B.A., San Diego State University

M.S., Ph.D., University of Massachusetts, Amherst

Directed by: Professor Alexander Pollatsek

How do people's theories of uncertainty differ from formal theories of probability? As an introduction to an investigation of this question, brief reviews were provided of a) various schools of thought within the science of probability, and b) research which has tried to determine how consistent people's behavior is with that prescribed by formal theories. While much of the research has suggested that, in many respects, people are poor intuitive statisticians, little has been said of how people actually arrive at probabilistic judgments. More recent research which has attempted to describe the strategies or heuristics that are used in reasoning about uncertainty have inferred reasoning processes from group performance on questionnaire items.

This investigation employed a more direct methodology: Sixteen undergraduates were interviewed and instructed to "think aloud" as they solved several word problems which required probabilistic judgments. Videotapes of the interviews were analyzed, and a model of

probabilistic reasoning was developed and described as an "outcome approach" to uncertainty. According to the outcome approach, the goal in questions of uncertainty is to predict the outcome of an event. Since the primary focus is on an individual trial (as opposed to a sample), predictions take the form of yes or no decisions of whether an outcome will occur on a particular trial. These predictions are then evaluated, after-the-fact, as having been either right or wrong. Moreover, rather than employing a chance or "black-box" model of uncertainty, outcome-oriented individuals often arrive at predictions by identifying factors that are believed to cause or inhibit certain outcomes.

The validity of the outcome approach was supported with correlational evidence based on coded portions of the interviews and with reportage of interview segments. In addition, predictions were made of how outcome-oriented subjects would respond to a different set of questions. These predictions were verified in a second set of interviews with 12 of the original subjects.

While the outcome approach was described as being inconsistent in several respects with formal theories of probability, it was portrayed as being internally consistent and valid in the context of everyday decision-making.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Probability is a particularly slippery concept. The difficulty has not so much to do with the fact that it is not one but a family of concepts made up of separate yet not quite distinct members; many concepts share this distinction. Nor is the difficulty a result of the concept's conversion by mathematicians and philosophers into a technical term; again, many of our concepts have been born again in scientific garb and caused much less of a stir. What is uniquely troublesome about probability is the territory it stakes out. It attempts to demarcate the amorphous state somewhere between total ignorance and perfect knowledge. And it is trying to keep one's footing in this nowhere land that is particularly disturbing. Like a frictionless surface, it not only trips you up, but keeps you sliding once you're down.

According to Hacking (1975), prior to the 17th century such a middle ground between belief and knowledge did not exist, yet the term probability had been around for some time. Hacking's thesis is too complex to review here in detail. Briefly, he argues that in the early to mid 1600's the concept of evidence changed such that information about the nature of things could, for the first time, be found in the things themselves. Prior to this, there existed a class of phenomena that could be known through demonstration (scientia) and another class of things that could only be testified to either by men of authority or through God-given signs (opinio). The word probability attached to the latter class such that an opinion was probable if it had received the stamp of approval of some authority figure. A "probable" belief was thus an "approved" belief. But no amount of approval could ever make an approved belief a matter of demonstrable knowledge.

Through the concept of "internal evidence," as Hacking calls it, a bridge was created between belief and knowledge. The old term, "probability," took on the function of describing the distance a belief had traveled from total ignorance to certain knowledge. At the same time, however, it also began to be used to refer to the tendency of certain phenomena (chance processes) to produce stable frequencies. This dual usage was no accident, for frequency data were a special variety of the new type of evidence that permitted degrees of belief separate from opinion. Now whether a belief was probable had less to do with approval of authority than with the "approval" of data.

If we were to meet a 16th century ancestor, we would no doubt get bogged down trying to understand what one another meant when we said that something was probable. Unfortunantly, we need not do any time traveling to observe or engage in a confusing dialogue about probability: We need but walk into a classroom where statistics or probability are being "taught" to find two parties floundering in an attempt to understand one another. If Hacking is correct in his analysis, a conversation with an early ancestor would be less troublesome because of our having understood, through Hacking, their concept of evidence and what it meant to "know" verses to "believe." My hope is that a similar understanding of how people informally view uncertainty may facilitate communication between teachers and students of probability. This thesis was undertaken with this goal in mind.

Some understanding of the source and nature of conflicting views of probability can be gained from an examination of the various theoretical schools of probability. Though they quibble amongst themselves over the issue of the appropriate domain of probability, in those areas where they overlap there is usually agreement concerning the numerical probability assigned to some event. Closer to the issue being explored in this study are empirical studies which have compared numeric answers given by groups of subjects to the ideal or normative values according to formal probability theory.

In the remainder of this chapter I will first provide a brief description of various schools of probability and then summarize the research which has attempted to determine how good the common person is at statistics. As a final introduction to a report of two original studies, I will stress the importance not only of generating data showing that the "informal probabilist" is non-normative, but of creating theories of how people reason about uncertainty--of why they arrive at the answers they do. The need for theories of this nature is especially critical if one's goal is to foster better communication in the teaching and learning process.

Facets of Probability

Our present concept of probability took form around 1660 when Pascal, Huygens, Leibniz, Fermat, and others somewhat independently developed formalizations for treating such diverse phenomena as games of chance, legal decisions and annuities. In a letter to Fermat, dated October 27, 1654, Pascal reviewed their independently-arrived-at solutions to a problem of how to divide stakes in an interrupted game of chance. He demonstrated that while, on the surface, their approaches appeared different, they were, in fact, comparable. He concluded with a statement that, in retrospect, seems prophetic: "Now our harmony has begun again" (Maistrov, 1974, p. 39). The harmony was two-part, consisting of the acknowledged need to discriminate between independent and dependent events, and of the addition and multiplication rules which allowed for the calculation of the union and intersection of independent events.

What came into being at that time was, to strain the metaphor, a new tune. What was not so new were the lyrics--"How probable is it that..." Since that time we have been plagued with a concept of probability which has at least two different but related aspects. On the one hand it is a somewhat vague epistemological notion of the degree to which some belief is held to be true; on the other, it refers, in a precise way, to the relatively stable frequencies exhibited by some phenomena. The difference between these two can be illustrated by comparing the two propositions below:

1) "My confidence is X that the statement, 'I will live to the age of 60,' is true."

2) "Accoring to life expectancy tables, my chances of living to the age of 60 are X."

The first is a combination of two propositions--"I will live to the age of 60" and, "My confidence is X that the statement is true." The nature of the expressed belief is not specific, but we can assume that it derives, legitimately or not, from a variety of sources. This is termed epistemological probability.

The second proposition is different in that there is no question as to the source of the number. It is derived from available frequency data. This is aleatory probability (cf. Hacking, 1975).

Not only have these two meanings clung tenaciously to the same word, but, since the time of Pascal, several schools have sprung up on both sides of the probability street. On the aleatory side, three schools have developed:

<u>Classical Theory</u>

According to the classical definition, which was prevalent until the time of Bernoulli, the probability of an outcome is simply the ratio of the number of favorable possibilities to the total number of basic alternatives. Thus, the probability of rolling a 3 on a 6-sided die is 1/6 since only one possibility is favorable to the outcome, while 5 are not. An obvious assumption, and the major limitation of this definition, is that the outcome alternatives must be equally likely.

Frequentist Theory

According to a frequentist theory (e.g., Reichenbach, 1949; von Mises, 1957) the probability of an event is its limiting frequency of occurrence in an infinite series of trials. Thus the probability of rolling a 3 on a particular die is the relative number of occurrences if the die were rolled an infinite number of times.

Propensity Theory

To the frequentist, the probability of an event refers to its frequency in the long run. One cannot, therefore, sensibly talk of the probability of a single occurrence. According to propensity theorists (e.g., Hacking, 1965; Mellor, 1971), however, the probability of rolling a 3 is an inherent, physical characteristic of the die. As a result, probability is viewed as something that belongs to each individual trial.

On the epistemological side of probability, two major schools can be distinguished:

Personalist Theory

According to this view (e.g., Ramsey, 1931; Savage, 1954), all probability is a matter of personal judgment, and is thus subjective.

By assuming that such judgments are internally consistent, however, a measurement technique is hypothesized that involves an exhaustive ordinal comparison of a person's judgments. Using Bayes' theorem as a normative algorithm for adjusting these judgments given new information, the theory can be used as a model of rational decision-making.

Logical Relations Theory

According to these theorists (e.g., Carnap, 1950; Jeffreys, 1948), the confidence in a proposition is the extent to which it is logically implied by some other proposition. Since the relations between propositions of this type are defined formally and do not rely on the specific content of the propositions, the theory is viewed as an objective as opposed to subjective theory. (For an in depth treatment of these various theories, see Mackie, 1973.)

While valiant attempts have been made by advocates of both aleatory and epistemological probability to annex the concept, the division has remained. Though arguments among the theoretical schools have generally revolved around questions of the internal coherence of the competing theories, the question of psychological validity has also been raised as a concern. The match between a probabilistic theory and actual behavior has been especially important in personalist theory. Similarity to actual decision-making has been cited as one of its strong points:

> The notion of probability which we have described is without doubt the closest to that of 'the man in the street'; better yet, it is that which he applies every day in practical judgments.... What more adequate meaning could be discovered of a notion? (de Finetti, 1964, p. 111)

Dempster (1968) points out that the entanglement of theoretical and behavioral considerations has its beginnings in the early history of probability:

> The earliest mathematical developments of the theory of probability arose from games of chance, where the central problem was to settle a rule of behavior, and the numerical probabilities implicit in the developed models were indentified after the rules of behavior. (p. 64)

In the late 1950's, psychologists interested in human decision-making began viewing statistical theory as providing a normative solution for judgments under uncertainty. These investigators wanted to determine the extent to which actual behavior conformed to the normative statistical solutions.

Peterson and Beach (1967), in their review of the research, concluded that probability and statistical theory provide a good first approximation of actual human decision-making under uncertainty. Subjects seem to be quite good at estimating descriptive parameters such as means and variances, and somewhat less able in tasks involving statistical inference. For example, subjects' solutions have been found to differ from normative solutions in the following cases:

1) While subjects use new information to revise prior probabilities, their posterior values are less affected by the new information than they ought to be (Edwards, Lindman, & Phillips, 1965).

2) In inferring parameters from samples, subjects tend to minimize the contribution of extreme cases (Beach & Scopp, 1967).

3) In making inferences about correlations from 2 X 2 contingency tables, subjects frequently rely on either one cell or on both positive diagonals (Jenkins & Ward, 1965; Ward & Jenkins, 1965).

4) Subjects underestimate the importance of sample size in assessing the diagnostic value of a derived statistic (Edwards & Kramer, 1963).

5) Over independent trials, subjects behave as if the probability of an outcome changes (Peterson & Beach, 1967). They also construct "random" sequences that have too few long runs (Brown, 1964).

While research comparing actual to normative judgments under conditions of uncertainty had produced some interesting findings, Peterson and Beach (1967) concluded their review by stressing the need for a more descriptive approach and the contruction of psychological theories that would better account for human probabilistic reasoning. Since that time, Daniel Kahneman and Amos Tversky have been the most prolific in generating explanations that account for non-normative statistical behavior. In a series of articles, they have suggested that because of limited information-processing capabilities, people use various judgment heuristics or rules that allow them to summarize large amounts of data and quickly arrive at decisions. These heuristics are described below:

1) With the <u>representativeness heurtistic</u>, the probability of a sample is assessed by noting the degree to which it shares essential characteristics with the parent population.

2) With the availability heuristic, probability is assessed according to the ease with which relevant instances of an event can be brought to mind.

3) With the adjusting and anchoring heuristic, new information is used to revise prior probabilities in such a way that the resultant values are closer to the initial probabilities. (Kahneman & Tversky, 1972, 1973; Tversky & Kahneman, 1973.)

While these heuristics usually produce adequate estimates, limitations in the amount and type of information that they are sensitive to leads to predictable judgmental errors in some situations. These heuristics, for example, have been used to explain the gambler's fallacy, insensitivity to sample size, and misconceptions involving random sequences and regression effects (Tversky & Kahneman, 1974). That these heuristics have been integrated into attribution theory and inspired some new lines of investigation in the field of social psychology is evidence of their generalizability (cf. Nisbett & Ross, 1980).

However, in terms of describing an informal theory of probability that would be useful in instruction, these studies have three major limitations. First, they have employed mostly questionnaire methodologies which provide very indirect information about how subjects are actually reasoning about probabilistic situations. Second, only measures of group performance are generally reported, which masks individual differences that could be especially important in instruction. Third, the judgment heuristics are not alternative theories of probability. Rather they are viewed as short-cut procedures for arriving at probability judgments. The question of the underlying theories people might hold is left unaddressed, yet the assumption inherent in asking subjects to provide numeric probabilities in response to questions is that subjects view probability consistent with formal theories. If this is not the case, subject responses cannot be easily interpreted.

More recently Kahneman and Tversky (1982) have attempted to describe several variants of probability as suggested in the usage of the term in natural language. Similar attempts have been previously made (Lucas, 1970; Mackie, 1973; Toulmin, 1958). The belief that informal theories of probability can be inferred from an analysis of language is in the spirit of Hacking's (1975) analysis and is a major assumption in the method to be used in this investigation of informal conceptions of probability. A brief review of the objectives and assumptions of this thesis follows.

Objectives and Assumptions

In this investigation I draw heavily on the image of probability as a concept that is "rooted in a dialogue" (Lucas, 1970, p. 9), that "exists in discourse and not in the minds of speakers" (Hacking, 1975, p. 16). Using the metaphor of a conceptual space much the same as Kuhn (1962) talks of paradigms or Foucault (1973) of epistemes, Hacking suggests that

> concepts are less subject to our decisions than a positivist would think, and that they play out their lives in, as it were, a space of their own.... All those who...employ the concept use it within this matrix of possiblities. (p. 15)

As previously mentioned, I view students and teachers of probability as possessing two different frameworks which, to the extent that they are compatable, make communication possible. To the extent, however, that they differ, communication is problematic, and persistently so. We can assume that the teacher's concept is similar to one of the schools (and probably to the frequentist theory) as summarized above. The major objective of this project is to explore and describe informal concepts of probability--the ideas held by non-theorists who, nevertheless, comes into contact with uncertainty daily; the ways by which they determine what "the chances are" in different situations; and the meaning of expressions related to such a determination.

The fact that the status of the concept of probability is disputed by theorists provides some clues as to the nature of the difficulties encountered in this study. No doubt the varieties of informal theories are staggeringly numerous if described in detail. I proceed on the assumption that the majority of people can be described as falling into a few categories. I assume, too, that peoples' notions of probability are, for the most part, internally consistent and valid in the context of their environment. By exposing them to a variety of problems, I have attempted to determine if and at what points their approaches to the problems change, and in this way gain access to their "informal probability axioms." This approach is ethnographic is spirit in that I attempt to provide phenomenal (world-as-perceived) as opposed to nomenal (world-de facto) descriptions, using research methods (the indepth interview) appropriate to such description (Magoon, 1977; Wilson, 1977; Marton, 1978). The basic design of the study involves an initial set of interviews with students who have had varying amounts of exposure to formal probability. From subjects' responses to several questions involving uncertainity, a model is hypothesized which accounts for their performance. From this model predictions are made of how subjects will respond to a different set of questions. These predictions are then tested on the same group of subjects in another interview. This last step is undertaken in an attempt to overcome the problem of the relatively subjective nature of interpretations of interview data (see Konold & Well, 1981).

My interest in this question has developed in two contexts. First, research on related topics (e.g., Well, Pollatsek & Konold, 1981) has suggested that we may have been underestimating the extent to which people view probability differently from theorists. I have been particularly impressed by subjects's unwillingness in interviews to provide quantitative estimates of probability. Second, the question has been motivated by attempts to teach probability in the context of a course in psychological statistics. It seems reasonable that to the extent that instructors have insight into the particulars of informal theories, they would be more effective in helping students understand alternative approaches. The differences between formal and informal theories may provide sources of explanation for the difficulties encountered in teaching; the similarities between the two approaches may point to inroads from one to the other.

CHAPTER II

METHOD: INTERVIEW 1

Subjects

Seventeen undergraduate students at the University of Massachusetts, Amherst, volunteered to participate in the study by placing their names on a sign-up sheet that was posted in the psychology building. Subjects were offered a choice of either extra course credit or $5.00 for their participation. They were informed that they would be asked to talk about what they thought would happen in various situations that involved uncertain outcomes, and that they would be asked to participate in a similar study the following semester. One interview was terminated before completion when it was thought that the subject was experiencing a high degree of stress. Data from this interview were not analysed. Information concerning the sex, major, and math and statistics background of each subject is included in Table 1, Appendix A.

Procedure

Setting

Subjects were interviewed individually. Upon entering the room, subjects were asked to read and sign a consent form which described the study as being motivated by the desire to better understand the kinds of relevant knowledge students possess before or after having completed a statistics course. They were instructed that they would be given several problems that would require them to predict what they thought would happen in situations involving uncertainty. They were told that the particular answers they gave were of less interest than their reasoning that led to the answer. Accordingly, they were instructed to "think aloud" as they attempted to solve each problem, verbalizing their thoughts as they occurred rather than attempting to reconstruct them at some later time. A felt pen and piece of paper were provided, and subjects were told that they could use these for any figuring or drawing that would help them to think or communicate their thoughts about the problems.

The subject and interviewer sat on the same side of a table and approximately three feet apart so that they would both appear on the videotape. Subjects were aware that they were being videotaped, and the recording equipment was in full view. I conducted all of the interviews. I had had no contact with any of the subjects prior to recruting them for the study. My experience with conducting and analyzing interviews of this type had begun about two and a half years before these data were collected.

Problems

Twelve problems were initially developed and piloted in twelve interviews with other subjects. These problems were chosen to vary on several dimensions that either were regarded as important within various theoretical perspectives (e.g., repeatability and independence of trials) or were thought of as potentially important to some individuals (e.g., past vs. future outcomes and outcome utility). Five problems were finally selected for inclusion in the first set of interviews. However, subjects' responses to only three of these problems will be treated in this report. These problems are presented below. (The organization of the problems into paragraphs is meant to indicate clusters of questions that were related and does not have any other meaning vis-a-vis the way questions were presented during the interview.) Standardized probes that were asked of most subjects are also included. Non-standardized probes for the most part consisted of requests to clarify or repeat a statement and reminders to verbalize. However, unique probes spontaneously developed to test hypotheses about subjects' reasoning were also employed. The interview format, therefore, could best be described as indepth (cf. Konold & Well, 1981) in which extensive probing by the interviewer is allowed, provided the probes do not direct subject thinking. Appendix B includes an entire transcript of the interview with Subject 11 and provides a representative example of the interviewing style and the nature of the two problems not included in this report.

Weather problem. What does it mean when a weather forecaster says that tomorrow there is a 70% chance of rain? What does the number, in this case the 70%, tell you? How do they arrive at a specific number?

Suppose the forecaster said that there was a 70% chance of rain tomorrow and, in fact, it didn't rain. What would you conclude about the statement that there was a 70% chance of rain?

Suppose you wanted to find out how good a particular forecaster's predictions were. You observed what happened on ten days for which a 70% chance of rain had been reported. On three of those ten days there was no rain. What would you conclude about the accuracy of this forecaster? If he had been perfectly accurate, what would have happened? What should have been predicted on the days it didn't rain? With what percent chance?

Numbers are often assigned to various events like elections, sporting events, causes of death. For example, you might hear something like, "A person's chances of dying before the age of 40 are 25%." What does the 25% tell you? Or you might hear, "The Equal Rights Ammendment has a 25% chance of being ratified by the June, 1982 deadline." What does the 25% in this case tell you?

Misfortune problem. I know a person to whom all of the following things happened on the same day. First, his son totalled the family car and was seriously injured. Next, he was late for work and nearly got fired. In the afternoon he got food poisoning at a fast-food restaurant. Then in the evening he got word that his father had died. How would you account for all these things happening on the same day? Have you ever had something very unlikely happen to you? To someone you know? How do you make sense of a day when everything seems to go wrong?

Bone problem. I have here a bone that has six surfaces. I've written the letters A through F, one on each surface. (Subject is handed the bone which is labeled A, B, C, and D on the surfaces around the long axis, and E and F on the two surfaces at the ends of the long axis.) If you were to roll that, which side do you think would most likely land upright? How likely is it that x will land upright? (Subject is asked to roll the bone to see what happens.) What do you conclude about your prediction? What do you conclude having rolled the bone once? Would rolling the bone more times help you conclude which side is most likely to land upright?

(Subject is asked to roll the bone as many times as desired.) What do you conclude having rolled the bone several times? How many times would you want to roll the bone before you were absolutely confident about which side is most likely to land upright?

One day I got ambitious and rolled the bone 1000 times and recorded the results. This is what I got. (Subject is handed the list which showed A-50, B-279, C-244, D-375, E-52, F-0.) What do you conclude looking at these? Would you be willing to conclude that D is more likely than B? That B is more likely than C? That E is more likely than A? If asked what the chance was of rolling a D, what would you say?

I'm going to ask you to roll the bone ten times, but before you do, to predict how many of each side you will get. How did you arrive at those specific values? (Subject rolls the bone and notes the results of each trial on the sheet of paper. After the 8th trial the subject is asked:) What is your best guess of what you will get on the next two rolls? (After the last trial the subject is asked:) How do you feel about your predictions? If you were going to roll the bone ten more times, what would you predict that you would get?

Presentation

The problems were presented orally by the interviewer. Two orders of presentation were used, the order being alternated on each successive interview. Order A was the sequence Weather, Bone and Misfortune followed by the two other problems. Order B was the reversed sequence.

Each interview required approximately one hour. Following the interview, subjects were told that questions about the project would be answered following the second interview which would be conducted during the next semester.

CHAPTER III

ANALYSIS AND DISCUSSION: INTERVIEW 1

The aim of this project is to articulate an informal approach to probability that allows peoples' statements and expressed beliefs concerning probability to be seen as interrelated and, ideally, logically consistent. The videotaped interviews and rough transcripts made from the videotapes provided a set of statements from which such a theory could be developed.

The analysis comprised two general and interrelated stages of hypothesis formation. The first stage involved repeated reviews of the transcripts and videotapes with the purpose of answering questions of the form, "What is this subject trying to say about x?" An answer to such a question is regarded as a descriptive hypothesis (cf. Konold & Well, 1981)--descriptive in the sense that it simply attempts to restate what the subject was attempting to communicate, and an hypothesis because we can never be sure that we understand what was meant to be communicated. Attention was <u>especially</u> focused on statements that seemed inconsistent with any formal theory of probability and that were common to several subjects.

The second stage was to attempt to infer from the collection of statements or descriptive hypotheses a set of underlying "axioms" from which the statements could be deduced. This stage is conceived by

Konold and Well (1981) as one involving the formulation of interpretive hypotheses--interpretive in that the attempt is made to go beyond what the subject is explicitly communicating and to infer what tacit knowledge the subject must possess in order that their expressions appear internally consistent.

In this chapter I will describe an informal approach to probability which was deduced from an analysis of the interviews and which I have labeled the "outcome approach." This approach will be decribed as one in which the primary objective is to predict the outcome of an individual trial. Given this objective, questions explicity asking for the probability of an outcome are translated into questions of whether, in fact, the outcome will occur. To give an initial impression of the approach, two composite interviews are juxtaposed below. On the left is a prototype of the outcome approach; on the right, a prototype of a frequency interpretation. These prototypes are assembleges of excerpts from several subjects (as noted) and should not, therefore, be regarded as typical responses. The prototypes are presented as ideal characterizations of subjects who reason according to the outcome approach on the one hand, and according to a frequency interpretation on the other. While a few of the subject protocols will closely resemble one or the other of these prototypes, most subjects, in fact, will be described as in a conflict stage somewhere between their previously-held outcome approach and a recently-exposed-to frequency interpretation. Many of the subjects interviewed had taken a statistics course in a university psychology department. Since the statistical approach of the social sciences is almost exclusively in the tradition of Fisher or Neyman-Pearson, subjects' exposure to formal probability had been frequentist in nature. For this reason, the conflicts that subjects are typically expressing are assumed to have resulted from attempts to incorporate principles and conclusions of frequentist theory with prior-held theories. If students who had been exposed to one of the other formal theories of probability were interviewed, different conflicts might have been manifest that could not be understood within the framework of the outcome approach. This represents a limitation of this study. However, since most of probability and statistics instruction is based on frequentist theory, the thinking displayed here should be typical of most students exposed to statistics.

The prototypes presented below will also provide some sense of the flow of the interviews. An entire transcript of Subject 11 is included in Appendix B and serves both to give a more extensive account of the nature of the interviews and as a good example of a subject caught in conflict between an outcome and a frequentist approach.

Table 2. Prototype of outcome-oriented and frequentist responses to the Weather, Misfortune, and Bone problems.

<u>Outcome Approach</u> | <u>Frequency Interpretation</u>

<u>Weather Problem</u>

I: What does it mean when a weather forecaster says that tomorrow there is a 70% chance of rain?

S6: I just think of that as a pretty good chance that there'll be rain. If he said 50% chance, I wouldn't think much about it. But when he says 70%, I would think of thinking about an umbrella.

S4: 70% means that the chances that it will rain are seven out of ten, according to him.

I: What does the number, in this case the 70%, tell you?

S12: Well, it's close to 100. It's more of a positive figure than, say, 30 would be. It represents a "C" in grades. Just a positive number.

S4: Well, it says that there's a 30% chance that it isn't going to rain.

I: Suppose the forecaster said there was a 70% chance of rain tomorrow and, in fact, it didn't rain the next day. What would you conclude about the statement that there was a 70% chance of rain?

S12: Well, that maybe they just fouled up. Or during the night, the precipitation or something changed in a different direction because of other outside factors.

S4: Well, on the basis of just the sample, I think an unrational response would be that the prediction was wrong. But, in fact, 30% is a pretty good probability that it's--it's not miniscule that it's not going to rain.

I: Suppose you wanted to find out how good a particular forecaster's predictions were. You observed what happened on ten days for which a 70% chance of rain had been reported. On three of those ten days there was no rain. What would you conclude about the accuracy of this forecaster?

S3: Well, I suppose he probably should do better than that. I assume they're trying

S2: He was exactly right. Seven out of ten times is 70%. And he concluded 70% chance of rain all ten their best. They're not trying to feed you wrong information.

times. So--70% of all the time.

Misfortune

I: I know of a person to whom all of the following things happened on the same day.... How would you account for all these things happening on the same day?

S11: Each of those is a separate event, and why they would happen on the same day, I don't know. I'm trying to look for clues, you know, trying to figure out why. I would probably just look at each one individually --if his son was just a bad driver or if he was a good driver and he'd never gotten into an accident--try to figure out something about-- Maybe he'd had it good for too long and there was just like all these bad things piling up, just waiting to come at him.

S2: It's arbitrary, somewhat. It just occurred. I don't see any other way I could explain how they all occurred on the same day. I could see how if the guy totalled his car, he'd probably be late for work. Even though it's unlikely to occur, like if it only happens 1 in 1000 times, if you live 1000 days the odds are it's going to happen to you. So even though it's unlikely for an every-day occurrence, when you consider all the days that you live, it's not so unlikely.

Bone Problem

I: If you were to roll this, which side do you think would most likely land upright?

S9: Wow. If I were a math major this would be easy. B is nice and flat, so if D fell down, B would be up. I'd go with B.

S2: I don't think I could tell you without rolling it. This is not like a die, and I think that there is no way of me knowing personally without experimentation.
S4: I could only give my

best guess. I'd have to say B up.

I: And about how likely do you think B is to land upright?

S9: I wouldn't say it's much more likely. It depends on the roll, I think.

S4: I'll give a big bias to B. I'll say 33%.

I: So what do you conclude having rolled it once?

S10: Wrong again. [B] didn't come up.

S15: I don't conclude anything. Can I roll it again?

I: Would rolling it more times help you conclude which side, if any, was most likely to land upright?

S9: No. I don't know. I think it's difficult to decide which is more likely. I don't see how you really can, just by looking at it. That's my opinion.

S1: Oh definitely. I mean that's the only way I could tell for sure. I think the only way with a thing like this is to just keep rolling it and just record the results.

These prototypes were assembled and juxtaposed primarily to highlight three features of the outcome approach:

a) the tendency to interpret questions about the probability of an outcome as questions about the outcome of a <u>single</u> <u>trial</u>;

b) the use of probability values as a means of arriving at a yes or no <u>decision</u> which, after-the-fact, can be evaluated as having been either right or wrong; and

c) the reliance on <u>causal</u> as opposed to statistical or chance explanations.

In the remainder of this chapter a more formal treatment of these features of the outcome approach will be undertaken. First, the features will be elaborated, and excerpts from interviews will be used to demonstrate their form and pervasiveness. Then, further evidence of the basic role and interrelatedness of the features will be offered in the form of a) their co-occurrence in subject protocols, and b) their tendency, in contraposition to principles of frequentist theory, to produce conflict states in certain subjects. Finally, as an introduction to the analysis of a second set of interviews, predictions will be made as to how outcome-oriented subjects will respond to a novel series of questions. This last step is viewed as particularly important because it provides the opportunity to subject the outcome approach to a more critical analysis in which disconfirming evidence of a more objective nature is sought.

Characteristics of the Outcome Approach

Outcome-oriented subjects, when asked the probability of repeatable events, believe they are being asked about the results of a single trial. This tendency to focus on the outcome of a single trial will be referred to as the _single-trial_ feature. This contrasts with the frequency approach where the primary focus is a _sample_ of trials, and in which probability is conceptualized as the relative frequency of occurrence of an outcome in a large sample.

Closely related to the single-trial feature is the objective of the outcome approach which is to predict outcomes of trials in a yes-or-no fashion. This objective will be referred to as the decision feature and, again, contrasts with the frequency approach in which the objective is usually to predict the average value or percent of some outcome in a large number of trials. The predictive feature in the frequency approach does not involve a yes or no prediction as it does in the case of the outcome approach. Rather, the prediction takes the form of an estimate of the long-run frequency of occurrence of a particular outcome.

The outcome approach will be seen to be more similar to a personalist approach in that in both cases a probability value gauges the degree of belief that a particular outcome will occur on a given trial. However, a personalist approach does not necessarily involve yes/no predictions. A personalist, having made the statement, "The probability of x is *p*," would attempt to verify the accuracy of this proposition by determining what percent of events assessed the probability *p* actually occur. For example, if a personalist estimated the probability of several independent and unique events to be .60, and these were "good" estimates, it would be expected that approximately 60% of these outcomes would occur, while 40% would not. In the outcome approach, a prediction is only "good" if the predicted outcome occurs: A prediction has no specfic meaning in relation to future or past predictions of the same or similar outcomes.

The remainder of this section is divided into two major subsections. In the first, the single-trial and decision features will be considered together and illustrated using excerpts from subject interviews. In the second subsection, the third feature of the outcome approach will be similarly elaborated and exemplified. This characteristic, referred to as the causal feature, involves the belief that chance phenomena can be analyzed in a mechanistic or causal framework. This will be contrasted with formal approaches to probability which are based on a chance or "black-box" conceptualization of uncertainty.

Predictability of Individual Trials

As mentioned above, in this subsection the single-trial and decision features are considered together. An explanation for why they are not being discussed separately is necessary for understanding coding procedures which will be described later and also will help clarify the perceived relation between these two features.

Though aspects of subjects' reasoning can separately be described as consistent with either the single-trial or decision feature, in fact, in these protocols evidence for these two features almost always occurs together in the same sentence or proposition. Thus, they appear as a functional unit such that outcome-oriented individuals see as their task predicting in yes/no fashion the results of a particular trial. This co-occurrence may be explained in part by the fact that the attempt was not directly made in the first set of interviews to explore how a subject's view of a series of trials might differ from that of a single trial. (Some data relevant to this question are presented in the discussion of the second set of interviews.) But there is good reason to believe that the virtually perfect co-occurrence of statements which support the single-trial and decision feature reflects a logical connection between the two: If the single-trial is viewed as the appropriate unit in the treatment of uncertainty, predictions which involve average number of outcomes have no basis. Average outcomes can only be computed and used to describe the probability of a single trial when a series of trials is seen as a legitimate unit of analysis over which an average can be computed.

Given a single trial as the unit of analysis, it therefore makes sense that associated predictions would simply involve a decision of whether or not a particular outcome would occur. The connection between these two features becomes even more understandable upon consideration of the fact that the majority of uncertain situations that people face are situations which will never be exactly repeated (cf. Fhaner, 1977) and for which an unequivocal decision of yes or no must be made (e.g., the question faced by a jurer of a defendant's guilt or innocence).

In the outcome approach, predictions of individual trials thus take the form of "yes," "no," and occasionally "I-don't-know" decisions of whether or not a particular outcome will be observed. Once a trial has been conducted, a yes/no prediction is evaluated as either "right" or "wrong." This evaluation of outcome predictions will be described first. Related excerpts from both the Weather and Bone problems will be presented to illustrate this tendency to evaluate a prediction as either right or wrong after a single trial.

<u>Retrospective evaluations of predictions</u>. In the Weather problem, a situation was posed where a prediction of 70% chance of rain had been made for a day on which no rain fell. Asked what they would conclude about the accuracy of the statement that there was a 70% chance of rain, six of the subjects responded that the statement must have been incorrect:

> S3: Well, he's inaccurate, but everyone's entitled to a mistake. I still wouldn't not listen to him just because he made a mistake that day.

> S13: Well, like I said, I don't really believe too much of what they say 'cause it's not always true. They said they thought it was going to rain, but it didn't.

> S11: ...and even if it was 90%--I would be carrying with me the idea that he...predicted that it was going to rain the next day with a 70% chance and he was wrong.... And now I'll take that into consideration on the next occasion, that they were wrong on one occasion.

Two of the subjects (one of whose response is included in the prototype above) suggested that when the prediction was initially made it was correct, but that things had changed since then:

> S9: Well, that it probably was correct at the time that he said it, but by the time it went out it probably wasn't really correct.

These evaluations of single-trial predictions as either right or wrong are inconsistent with a frequency interpretation which is concerned with a sequence of trials rather than with the occurrence or non-occurrence of a single predicted outcome. In the context of the outcome approach, however, a yes or no prediction of a single trial is the objective. And given a yes or no prediction, it is appropriate to evaluate the prediction as either right or wrong once the outcome of the trial is known.

Subjects' performance on the Bone problem provides further evidence of the pervasiveness of this tendency to evaluate outcome predictions as right or wrong. In the Bone problem nine of the subjects made comments to the effect that their prediction of which side would most likely land up was either right or wrong. In the case of Subjects 1, 3, and 9 these statements were uttered spontaneously after they had rolled the bone the first time:

S1: It was C. All right. Well, I was wrong.

S9: So it was the other surface. Well, it's kind of similar to the other sides I said. I guess I just took a guess--a gamble. I guess it could land on any of them.

At the end of the Bone problem, Subject 3 had concluded that B was the most likely, and as the bone was being removed, she interjected:

S3: You mean I can't roll it to find out?

I: Oh, sorry. Go ahead and roll it.

S3: Watch me blow everything I said. [Rolls a B] B. Oh, I was right.

The remainder of the evaluative statements made by Subjects 5, 10, 11, 12, 13, and 16 came in response to the probe, "What do you conclude from having rolled the bone once?"

S5: That I was right. That I logically worked out which would be the most likely one to come up and I was right.

S11: I guess it was right. So I guess I looked at the right things, or whatever.

Most of the subjects' comments seemed to apply to both the results of that roll and their prediction. However, more consistent with a frequency interpretation, two of the subjects, while declaring that they had been wrong, suggested that perhaps their prediction was still right but had just not occurred on that particular roll:

S16: I only rolled it once. The one time I rolled it I was wrong, but it doesn't mean that it's incorrect.

S10: Wrong again.

I: Why do you say you were wrong?

S10: Well, D didn't come up.

I: What do you conclude from having rolled it once? Do you question now whether D is the most likely one to come up or--?

S10: Well, I'd question it a little, but that was only one roll.

Evidence from both the Bone and Weather problem has been offered (and is summarized in Table 3, Appendix A) which supports the claim that subjects take questions explicitly requesting probabilities and encode them into ones that ask instead for a decision of which alternative will occur on a particular trial. Once the trial has been conducted, these predictions can retrospectively be evaluated as either right or wrong. Performance on the Weather problem also provides the opportunity to explore the nature of the actual predictions.

Yes/no predictions. In the Weather problem, subjects were first asked what it meant when a weather forecaster predicts a 70% chance of rain. Answers to this question provide an indication of the nature of predictions in the outcome approach. In the following excerpts, rather than interpreting 70% as the probability of rain, subjects apparently encode 70% as "rain, but I could be wrong," often using the values 100% ("definite rain") and 50% ("I don't know") as extremes with which 70% can be compared.

I: What does the specific number, in this case the 70%, tell you?

S12: Actually, I never really thought about it. 70%. Well, it's close to 100. It's more of a positive figure then, say, 30 would be. It represents a "C" in grades, which students tend to think about a lot. Just, you know, a positive number.

S3: Well, 70%--more than likely it's going to rain. I think 50%--it's 50% it will, 50% it won't. But 70% means more than likely it will rain.

S13: That there is more of a chance of it raining than not raining. Like if he said, maybe, 55% chance, then I'd be--well 50/50--even if he said 50% chance, I still think I'd stick more to it being a nice day out then it raining. But 70% pushes it a little bit closer to definite rain.

S15: ...And when I hear a probability, like the 70% chance, I think, "Well, it's going to rain," you know. Because it's a higher probability I take that as fact.

S6: I just think of that as a pretty good chance that there'll be rain. If he said, like 50% chance, I wouldn't think that much about it. But when he says 70%, I would think of thinking about an umbrella, or something like that.

I: What does the specific number, in this case the 70%, tell you?

S6: Well, it tells me that it's over 50%, and so, that's the first thing I think of. And, well, I think of the half way mark between 50% and say 100% to be like, well, 75%. And it's almost that, and I think that's a pretty good chance that there'll be rain.

Two observations are worth noting in regard to the above excerpts. First, it is clear that in the above protocols, subjects translate the statement "70% chance of rain" into the more definitive and qualitative statement, "It's going to rain." Secondly, this translation is accomplished by using the probability range of 0% to 100% to form a decision dimension, with 100% corresponding to "yes," 0% to "no," and 50% to "I don't know." Intermediate values are ultimately associated with one of these three anchor or decision points according to a vague and variable proximity criterion. Thus, 70% is considered sufficiently above 50% to warrant identification with 100% or "yes" with perhaps some associated expectation of error.

Given this qualitative (yes/no) interpretation of the probability range, 50% is not viewed as a predictive forecast, but as an admission by the forecaster of total ignorance about the outcome.

S14: ...Anyone can say 50% chance just by looking up in the sky. If it's cloudy, you can say there's a 50% chance of rain.

S9: Well, it's not 100% chance, and it's not 50/50, so he's not guessing. I don't know. If he said 50/50 chance I'd think that was kind of strange.

I: What would it mean if he said 50/50?

S9: I guess I would think that he didn't really know what he was talking about, because only 50/50--"it might rain or it might be sunny, I really don't know," you know. It's kind of weird.

I: So when he says 70%--

S9: 70--well. Sounds like he's saying, "The likelihood is that it's going to be rain, but then it might be--the clouds might clear up tomorrow or something like that, so it might change."

In Table 3, responses are noted which are indicative of subjects' having translated probabilities into qualitative decisions. Using probabilities to arrive at decisions in this way, however, does not necessarily imply that the meaning of the values as probabilities is not also available. A frequentist or personalist might hear the statement, "There is a 70% chance of rain today," and conclude that they had better take an umbrella. But they would know that if they looked at the record of those days for which a similar prediction was made, on approximately 70% of them it would have rained. One of the probes used in the Weather problem inquired about the accuracy of a forecaster who had predicted 70% chance of rain for ten days, when in fact, no rain was recorded on three of the ten days. According to probability theory, the proportion of rain to no-rain days is perfectly consistent with the 70% forecast that was made on each day. Three of the subjects' responses were consistent with this reasoning. Nine of the subjects, however, felt that the forecaster was only "pretty accurate," suggesting that there was room for improvement. Two examples are provided below.

> S10: Well, out of the ten days there was 70%--seven of them it did rain, three of them it didn't. I would say he was pretty accurate. I don't know about perfect.

> S12: I'd say it was pretty good--only three misses out of ten days. I would tend to trust his judgment more than if someone missed seven out of ten--got three right.

These subjects seemed to regard the days it rained as "hits" and the days it didn't rain as "misses." Since the forecaster was only right on seven of the days rather than on all ten, he was only "pretty accurate." Presumably this is the case because in the outcome approach the 70% has meaning only in relation to the particular day for which the forecast was made, and the objective is to predict the outcome on that day.

All but one of the nine subjects were also asked what should have been predicted on the days it didn't rain. In every case, subjects responded that the percentage given should have been lower than the 70%, and usually lower than 50%. This implies that on the days it didn't rain, the prediction of a 70% chance of rain was incorrect. Thus, this is an additional instance of evaluating a prediction as either right or wrong once the results of the trial are known.

S10: Well, out of the ten days there was 70%--seven of them it did rain, three of them it didn't. I would say he was pretty accurate. I don't know about perfect.

I: ....What should he have predicted on the days it didn't rain?

S10: Probably 30% chance or less that it would rain.

S12: Well, he could either have said that there's a chance that it might rain rather than being more definite, or just said "mild," you know, "some clouds," or something like that rather than being specific.

S15: That he had a high probability of predicting rain, because out of ten predictions, only three went to the low odds.

I: What would have happened if he had been perfectly accurate?

S15: Well, then he'd be really--chances he'd be really good.

I: ....What should he have predicted on the days it didn't rain?

S15: A lower probability of rain. Instead of saying 70%, 50%--or whatever. Maybe he did the probabilities wrong or maybe that's just the way it happened.

It would appear that for these subjects, the meaning of 70% as a probability is completely dominated by its use as a value to arrive at a decision of rain.

Evidence for the decision and single-trial features has been offered in this section without attempting to separate subject statements which are suggestive of one feature and not of the other. This is because yes/no predictions in every case have referred to single trials. That the single-trial feature is a separate component, however, is suggested by subject responses to the probe inquiring about the accuracy of a forecasters' predictions over ten trials. Responses to the effect that the forecaster was pretty accurate suggest that these subjects are evaluating accuracy by focusing on the results of each individual trial, then summing over trials to get an overall level of accuracy. In this case perfect accuracy would entail correct predictions on every day. Two of the subjects' protocols offer strong support for this interpretation. In attempting to resolve their own conflicting views as to whether the forecaster was perfect or pretty acccurate, they focused on the question of whether the individual day, or all ten days, was the appropriate unit of analysis.

> S8: Well, he's looking at an individual day--particular day--and he's setting up a percentage on one day. And you can't really extend that to an amount of time, I don't think.

> S11: ...they didn't predict a ten-day span; they predicted each individual day. So it seems as though when each day comes up it's a whole in itself, and it's not necessarily put together in a unit.

In this section the outcome approach has been described as being motivated by the desire to predict outcomes. This is in contrast to the objective of interpretations based on a formal theory of probability. Formal theories provide a system for estimating the probability that some outcome obtains given some chance set-up. Predictions for individual trials or a series of trials may be viewed as logically flowing from these probabilities. However, in most cases the outcome of an individual trial provides little information concerning the accuracy of the probability estimate on which it was based. In the outcome approach, one determines whether or not an outcome will occur and associated numeric values gauge a person's confidence in the prediction. If the outcome occurs as predicted, the predictor was right; if not, the predictor was wrong and the numerically expressed confidence was incorrectly arrived at. When probabilities are provided, as in the Weather problem, the outcome-oriented individual does not interpret them as probabilities per se, but as values which can be used to partition a decision space into yes, no, or I-don't-know decision regions.

Predicting Outcomes From Causes

In this subsection it will be argued that outcome-oriented individuals arrive at predictions through consideration of factors that would cause or bring about an outcome. Thus, the outcome approach will be portrayed as involving a causal as opposed to a chance interpretation of uncertain events, the latter being characteristic of formal probabilistic reasoning. A few comments about how uncertainty is viewed within a formal framework will help clarify the distinction between causal and chance explanations.

First, it needs to be stressed that a formal probabilistic approach does not necessitate the denial of underlying causal mechanisms in the case of chance events. Hypothetically, one can imagine describing the last in a series of 100 tosses of a fair coin in sufficient detail such that it could be seen to be determined by events which preceeded it. Certainly, its position on the back of a hand, say, was determined by its orientation in the hand in which it landed, which was determined by its position in the air the instant before landing, which ultimately was determined by the way in which the toss was executed, which, in turn, was determined by events including the previous toss, and so on. In this analysis, successive trials would be seen not as independent, but as dependent elements in a causal sequence. And as Venn (1962) states, a theory of probability "makes no assumptions whatever about the way in which events are brought about, whether by causation or without it" (p. 236).

In practice, however, a causal description is often seen as impractical if not impossible:

> It is a pure illusion to think that the motion of the balls on Galton's Board can be given a 'causal' explanation by means of the differential equations of classical mechanics, because...these equations can be said to determine unambiguously the course taken by a ball. And if we try to arrive at a better understanding of the motions by applying more and more exact experimental methods, we shall soon find ourselves in new difficulties.... Instead of finding something which...is more simple than the original phenomenon, we arrive at phenomena that are more and more complex. (von Mises, 1957, p. 208-209)

Accepting a current state of limited knowledge, a probabilistic approach adopts a "black-box" model according to which underlying causal mechanisms, if not denied, are ignored. Abandoning a specific, mechanistic analysis of cause and effect, a chance model gains its power from observations related to the fact that regularities occur over series of trials independent of specific causal agents. As will be illustrated in this subsection, a mechanistic model is not abandoned in the outcome approach. The goal of predicting the results of individual trials in a yes/no fashion would, in fact, seem to imply the possibility of determining beforehand the results of each individual trial. This could only be accomplished through an analysis of the appropriate causal mechanisms.

I will first discuss the Weather problem in which some subjects believe that the probability of rain is arrived at directly by measuring factors associated causally with rain, such as humidity, barometric pressure, and cloud cover. Next I will discuss the Bone problem in which several subjects apparently generated predictions and associated confidence values by attempting to analyze the features of the bone and, in cases, the rolling technique which would bring about (cause) different outcomes. Finally I will discuss the Misfortune problem in which subjects asked to account for the co-occurrence of several low-frequency events relied on causal explanations that involve, for example, imbedding the events into a causal-linear schema where each event but the first can be seen as a direct result of a preceding event.

Weather problem. In the Weather problem subjects were asked to explain what the specific number in the proposition "70% chance of rain tomorrow" means. Three types of interpretations were given by subjects. Three of the subjects (1, 11, 14) suggested that the number may be a measure of the confidence of the forecaster that it will rain tomorrow, though none of these subjects was clear about how the specific number was chosen.

I: What does the specific number, in this case the 70%, tell you?

S1: It would tell me that the weather forecaster thinks it's going to rain, but isn't certain. Because, say I got a test back and I got a 70. I knew enough to get a C but not enough to really do well.... I guess they just look at the kind of weather drifting into the area, but there's always a chance that it will veer off or something.

S11: ...just from conditions, I guess--from the atmosphere and from variables. I guess the weather is just like a lot of problems where you're looking at a whole bunch of variables and as they come together in one situation.... If you look out and see clouds here and they're so high up, and there's moisture here--then it kind of adds up to a likelihood that it will do something, and maybe the 70% is--maybe when they say that it has something to do with how sure they are that it's going to happen, I don't know.

Four of the subjects (2, 4, 5, 6) interpreted the 70% in a way that is consistent with a frequency interpretation, though as indicated in the excerpt from Subject 2, not necessarily in accord with accepted convention.

S2: It means that 70% of the area to which the broadcast is being given will experience precipitation.

I: Where did they get the specific number?

S4: I'd say a combination of current data and past records--just trends and any kind of history of the behavior of whatever data they have.

Three of the subjects (3, 9, 5, the latter having also given a frequentist account) suggested that the 70% was a measure of the presence of factors that would produce rain.

I: How do they arrive at that specific number?

S3: They must use their equipment to measure moisture and clouds. I don't know how it works at all, but it's fairly accurate--they must know what they're doing.

S5: What it means is they can see all these cloud patterns forming and moving into a particular area, but they're not as dense as, say, a hurricane where you can absolutely predict where it's going to go. 100%--that means it was a total cloud thing coming over the area.

S9: I guess he means that there's not a 100% chance, but the likelihood is that it's going to be a 70% chance of rain, I guess. Maybe it means that there's 70% humidity in the air, I don't know. I'd say the likelihood is that it might not rain that much, or something--just a little bit.

Though Subject 12 was not clear about what she thought the 70% specifically referred to, she responded to another probe which indicated that she may have a tendency to interpret probability values as measures of causal factors.

I: If you heard a statement like, "A person's chances of dying before the age of 40 are 25%," what would the 25% mean to you?

S12: Twenty five in that case would mean to me that the outside possibility, in other words, 25% would be the disease or whatever would happen to the person in order to make them ill or to the point where they would die, I think.

I: Can you say just a little more about that?

S12: I guess 25% would mean the consequences that would bring about someone dying before the age of 40.

I: Uh-huh. So 25% now reflects--What happens 25% of the time? Or 25%--

S12: Is the frequency of occurrence, I guess. Something that

> would bring about death--like a disease or maybe hereditary factors or something.

Interpreting the 70% as a measure of the presence of causal factors is consistent with regarding the occurrence of no rain as a mistake rather than as an event that had a 30% chance of occurring. Either the forecaster made an error in measuring or interpreting the factors or, as three subjects suggested, the factors were altered after they were measured.

> S8: I wouldn't think he'd goofed. I'd just think that for some reason or other, the atmosphere didn't--there wasn't a very predictive forecast. Weather changes a lot, you know.

> S9: Well, that it probably was correct at the time that he said it, but by the time it went out it probably wasn't really correct.

> S12: Well, that maybe they just fouled up. Or during the night the precipitation, or something, changed in a different direction, because of other outside factors.

<u>Bone problem</u>. Whereas the Weather problem provided the opportunity to see how subjects interpreted a probability expressed as a percentage, the Bone problem permitted the study of how subjects arrive at initial predictions under conditions of uncertainty and how they generate a number when asked how likely they think it is that their prediction will occur. It also provided the opportunity to explore how initial predictions are modified as new information is acquired. As in the Weather problem, excerpts from subject interviews which are consistent

with either a freqentist or outcome approach will be provided. Several subjects will be described as having turned to a causal analysis of the bone in attempting to answer the question. In this analysis, rather than relying on the results of previous trials, the prediction and associated confidence is obtained from an inspection of features of the bone, such as its surface area, which would tend to cause the bone to land one way as opposed to another. Before turning to a discussion of this tendency, however, some observations are in order concerning how subjects responded in general to the request to determine the side most likely to land upright.

Subjects were presented with the bone, but were told not to roll it before making an initial guess as to which side was most likely to land upright. Subjects, for the most part, seemed to regard this as a reasonable request. Two subjects, however, responded in a way characteristic of a frequentist, both feeling that they could not really say anything without first rolling the bone.

> S2: I don't think I could tell you without rolling it a few times.... This is not like a die, and I think that there is no way of me knowing personally without doing experimentation.

> S14: Predict? I couldn't 'cause I don't know which side is--It's like rolling a weighted die. You have to know the probability first. I don't know the probability.
>
> I: Why don't you go ahead and roll it and see what happens.
>
> S14: That doesn't matter, 'cause any of them could come up.

After subjects had made an initial prediction and had rolled the bone, they were asked if rolling the bone more times would be of further help in determining which side was most likely. For the frequentist, this question verges on silliness, the answer being all too obvious.

> S1: Oh definitely. I mean that's the only way I could tell for sure. I mean I could be totally wrong. I think the only way with a thing like this is to just keep rolling it and just record the results and somehow formulate which one would be the most likely to land on.

Consistent with this frequency interpretation, the majority of subjects used the information obtained from several rolls of the bone to modify their initial prediction such that the most frequently rolled side was believed to be the most likely. When shown results from 1000 trials, they also revised their conclusion, if necessary, and relied more heavily on the rolls from the 1000 trials then on their own observations when asked to predict the results of ten trials.

A number of responses, however, were inconsistent with a frequency interpretation. Five subjects (as noted in Table 4, Appendix A) expressed reservations about whether additional trials would be helpful in determining which side was most likely. Three subjects did not use the data provided from the results of 1000 trials in predicting the results of 10 trials. Eight subjects attributed variations among trials to the way the bone was rolled. Three protocols are reproduced below. These subjects seemed to believe that in answering the question concerning which side is most likely, more reliable information could be obtained from careful inspection of the bone than from rolling it, and that, in some cases, rolling it might even provide misinformation.

I: Would rolling it more help?

S6: Probably.

I: Why don't you roll it as may times as you like?

S6: [Rolls B,C,A] This isn't helping me any. [Rolls B]

I: When you say this isn't helping you any--?

S6: Just 'cause all those sides are coming up. [Rolls B,D,C] I didn't think that--[Rolls E] See, I didn't think that would ever show up, just because I didn't think it would stand like that. I'm surprized that it doesn't show B or D more [Rolls A] than it's showing. [Rolls A] I think I probably looked at the rock in a different way. I think the shape is kind of deceiving, so that I didn't realize it would stand up in so many different sides.... [While setting the bone on its various sides] I still think that, I don't know, it seems to stand the best on B and D really.

I: Do you have any idea of which side has come up most in the times you've rolled it?

S6: Well, B and D came up twice. A came up about three times, and C came up once. I don't think E or F ever came up. They don't really stand too well. [Actual results, including an initial D, were 3A, 3B, 2C, 2D, E] So I guess A is another side that comes up pretty often. But A and B are kind of--if it just rolled over once more it would go to B [Turns bone over from B to A]. So it would still make me think that B would be more prone to go than A, 'cause when it's showing up A it's not as sturdy as when it goes to B.

I: How many times would you want to roll that before you were absolutely confident about which side was most likely?

S6: Probably about 50 times. You could really calculate--even 100 I guess you need to really be sure. But still, I don't know. That doesn't seem to help me that much because it could just be, you know. If you rolled it 100 times it could show up on D every other time, and then the next 100 times you do it, it could be different results. So I don't know how much that would show either.

Heavy reliance on beliefs about the stability of various sides of the bone helped this subject maintain the correct belief that D or B were most likely to land upright despite the fact that the limited frequency data she had obtained suggested otherwise. This tendency is justifiable within a personalist theory of probability. However, inconsistent with any established interpretation, she also was uncertain that rolling the bone would eventually provide more reliable information than simple inspection of the bone. Subject 9, whose protocol follows, similarly did not believe that rolling the bone would provide the needed information, nor does she feel capable of properly analyzing the bone. Rather, she believes that to appropriately answer the question, the bone must be analyzed by someone more mathematically sophisticated than she.

I: If you were to roll that, which side do you think would most likely land upright?

S9: Wow. If I was a math major, this would be easy.

S9: ....Boy, this is hard. Yeah, I think D. But then again, B is very uneven which takes away from--But D is nice and flat, so if D fell down, B would be up. I'd go with B, I think.

I: OK. And about how likely do you think B is? B to come up?

S9: I couldn't even judge. I mean, my opinion is that it would be a little more likely than the others--just because the surface of D. But otherwise, I wouldn't say it's much more likely. It depends on the roll, I think. Probably.

S9: ....[Rolls C] C. So it was the the other surface. Well, it's kind of similar to the other sides I said.... So, I guess it could land on any of them. I don't know if you can determine something like that. Maybe you can with--mathematically. But I don't know.

I: How would you determine it mathematically?

S9: I guess you would because of the weights and stuff. However they do it. I know that some people have dice that are weighted down in certain numbers. And I suppose if they figured that out mathematically--which was more weighted--and figure surface structure.

I: ....Would rolling it more times help you decide which side was most likely?

S9: No. I don't know. I think it's difficult to decide which side is more likely. I don't see how you really can, just by looking at it. That's my opinion.

I: Why don't you roll it a few times and see what happens?

S9: [Rolls A,B,D,A,D,E,D,E,D,B,E,B]

I: ....What do you conclude having rolled it a few times?

S9: Well, I still think the possibilities are pretty even. I should have really counted out how many times each different surface hit.

I: That would help you?

S9: Probably. But...it seemed pretty even.... It wasn't more, one or the other, except for it was less E.

I: How many times would you have to roll that before you could be real confident about which, if any, was most likely?

S9: Oh, probably quite a few times. Like maybe 20.

After showing her the results of 1000 rolls she concluded that B, C and D were most likely. Then she was asked,

I: Would you conclude from looking at that [1000 results] that D is more likely than B?

S9: [Sets the bone with D, then B up, then looks at the results] Oh, yeah. Well, if the percentages are there, yep.

She used the data from 1000 rolls to predict the results of 10 rolls, but began to doubt the frequencies after she had rolled 2A, 1B, 1D, and 3C. After she had completed the ten rolls, rolling two more Ds and a B, she believed the statistics were "correct" but suggested that if it were rolled another 1000 times "that it might be totally different." If she were to roll it another ten times she said that she would use her previous results on the last 10 rolls in preference to the data from 1000 rolls.

The response of Subject 12 to the Bone problem provides the most extreme example of someone maintaining a belief based on the physical properties of the bone, despite frequency data which runs contrary to her belief.

I: If you were to roll that, which side do you think would land upright?

S12: ...I think B would be up.... D looks like it has four fairly flat corners. Then I could be totally wrong.

I: How likely is it that B will come up?

S12: I'd give it about a 75% chance.

S12: ....[Rolls a B] Ha. What do you know.

I: What do you conclude about your prediction?

S12: I guess my reasoning is somewhere on the ball. I think if you inspect things you can understand them, or if you question things, you can understand them better than if you just go ahead and make some free decision about something.

I: What do you conclude from having rolled it once?

S12: That my prediction was just about on the nose.

I: Would rolling it more help you decide which side was most likely?

S12: I don't know. I'm pretty sure about my prediction. But just for the fun of it I could do it--you know, just curiosity sake.

S12: ....[Rolls D] Uh-oh.

I: "Uh-oh" what?

S12: It landed on the opposite side.

I: So what does that make you think?

S12: That there's more than one possible answer to something--to this problem.

S12: ....[Rolls D, C] Oh. Well, I'm getting a little frustrated. 'Cause when you first get something right you're really happy about it.... But then when it turns out another way, well, that's OK too. But it starts to make you a little shaky. Now that this has happened--the third a C--well, I'm beginning to think this rock is quite equal all around.

I: So by "equal all around"--what do you mean?

S12: That it could be stable landing on any of the sides. It's just a matter of, sort of, chance, maybe.

I: ...Any idea of which side is most likely to come up? Or is there one that is more likely than the others?

S12: Well, I think D, C or B are more likely than E or F. Then I could be wrong.

I: That's on the basis of--How do you know that?

S12: Well, E and F have less of a mass around them. [Rolls C,D,B,B] Yeah. I think I'll go along with that last statement.

I: D, C or B?

S12: Yeah.

I: And do you think one of those--D, C or B--is more likely than the other two?

S12: Yeah. C.

I: Why?

S12: I don't know. The top part--it looks most stable because it's like a pyramid shape. The bottom is very thick and long.... It has all that base under it.

I: Would rolling that a lot more times help you, in any way, decide which side was most likely to come up--whether C was most likely?

S12: I don't think so.

S12: ....[Referring to data from 1000 rolls] So D you found most likely.

I: ....Would you feel safe in concluding looking at this, that D is, in fact, more likely than B?

S12: I don't know. Not really. I think it could just have something to do with your luck, or your chance. If I did the same number of rolls, I don't know if it would come out the the same.

I: ....What is your belief about which side is most likely to come up if you rolled it?

S12: I have to look at this again. It changed shape in the past few minutes. [Inspects bone] I still think C.

The above excerpts indicate, as previously suggested, that some subjects believe that the answer to the Bone problem is to be found in an inspection of those physical properties of the bone that would cause it to land in various orientations. The excerpts also indicate that the subjects are uncertain, to varying degrees, whether the results from previous rolls are useful in making inferences about the relevant dispositions of the bone. It might be that they distrust frequency data because they can seem so variable, whereas the characteristics of the bone that cause it to land as it does are viewed as invariable. This may explain why many subjects chose to attribute variability among trials to the specific way in which the bone was rolled rather than to view variability as a predictabile feature of a chance set-up. The "rolling hypothesis" was used by subjects both to explain differences between their data and the results of 1000 rolls, as indicated above, and also to explain variations in their own results. As an example of the latter type, Subject 5 had previously rolled B,B,D,C,D,C,D having initially predicted B.

> I: What are you thinking about?
>
> S5: I'm thinking that I threw it differently the first two times--that these times I'm throwing it from a higher height. [Rolls D] Now if I try it lower [Rolls D] it's still D. Now I wonder what I--[Rolls B] There's a B. [Rolls C] Now there's a C.

Another possible, and more basic, explanation for subjects' distrust of frequency data is that they may view them as a secondary and thus less reliable type of information when compared to the physical properties of the bone. Frequency data do not precede or cause the properties of the bone; rather, properties cause the frequency data. These subjects may feel that if one wants to predict what will happen on a next roll or series of rolls, they must predict with knowledge about those factors that would cause various outcomes. While frequency data reflect those causal factors, they are themselves not causal. Both the next outcome and the accumulation of past outcomes are products of the same causes--the properties of the bone and the way in which it is rolled.

<u>Misfortune problem</u>. Explaining the Misfortune problem by attributing the co-occurrence of the four events, or at least three of them, to chance is in accordance with a frequency interpretation. Eight of the subjects responded in this way. The protocol of Subject 2 in the prototype interview (Table 2) is a good example of the reasoning these

subjects provided. Five other subjects offered a chance account of the Misfortune problem as one possible explanation but suggested other explanations as well. Three general types of alternative explanations were offered: a) the attribution of "bad luck" to the person involved, b) the belief that some external force such as God or the stars caused the events, and c) the belief that the four events were causally connected such that one event triggered the next in the sequence of events. Subjects who offered two or more explanations typically did not state a preference among them. The exception was Subject 14 who preferred a chance to a causal-linear explanation. Table 4 summarizes the nature of the explanations that were offered by each subject. A few protocols illustrating the use of different types of explanations are provided below. (The protocol reproduced in the prototype interview is an example of both the casual-linear and external-force explanation. The latter was in the form of the belief that bad things pile up, eventually forcing themselves into one's life.)

Though several subjects had mentioned the possible explanation that the person had bad luck, they immediately dismissed it as implausable. Only Subject 3 seriously considered it as an explanation.

> S3: Has he always had bad luck or were things normal till this day pretty much?

> S3: ....I don't know why some people are lucky or some people are unlucky. That's the way it is though. And if he'd always had bad luck, you could almost expect something like that. And if he hasn't, if he'd been normal, then I don't know--weird.

Subject 5 showed a strong preference for establishing a causal-linear connection among the events, going so far as to reorder the events in a way that would make such a connection more feasable. He then seemed to dismiss a "cause-and-effect" account only to consider it again. He dismissed an external-force explanation and seems in the last sentence to have preferred his "cause and effect" account.

S5: Well, what caused what?

S5: ....I'm trying to figure out if the order you gave me was the order that they happened, or if his father died--or he went out to a restaurant with his parents--with his family to a fast food place, and they got food poisoning, and because he was sick, while he was driving he smashed up the car. His father died in the accident, and he was on his way to work so he was late.

I: [Repeated correct order of events]

S5: Things happen and there's no cause and effect going on which makes all these things happen at once. It just happened that way. There's no cosmic stars making it all happen. Just these things happen to people. His father may have died as a result of a heart attack from hearing the news that his son had a tragic accident....

S5: ...You can't find you future in the stars. I mean, it's very clear, if you turn left what's going to happen is left; if you turn right--voila. If you study physics all your life, you're going to become a physicist--Well, you could become a garbage collector, but the likelihood is--

Subject 13 seemed fairly confident that the events described in the Misfortune problem were a result of a divine force, though she does not feel that this is true of all events.

S13: I'm a strong believer in God--I'm very religious. So I just believe that when something happens like that...it just--like, "All right, this is the situation that's supposed to happen to me today. I got to handle it the best I can." But if it's something I know I screwed up on...then it's like it was my fault.

I: If someone said it was just coincidence, how would you feel about an explanation like that?

S13: I believe in coincidence, but not as much as I would believe that things are done for a reason, like, or things are done for a purpose. So I don't know if I'd say that that was a coincidence.

Though the three non-chance explanations differ in many respects, they all are types of causal explanations in the sense that they describe some condition that existed prior to one or more of the events and that produced and thus explains their co-occurrence. This tendency, to look to the antecedents of events both as a source of explanation and as a basis for prediction, is evident in all three of the interview problems. Predictions based on the frequency of past occurrence of an event are, in cases, consciously shunned.

Later I will discuss the functional value of causal interpretations of probabilistic phenomena and elaborate in more detail the difference between the frequentist and outcome-oriented views of chance events. What has been argued so far is that subjects have a strong tendency to look for and base predictions of future events on causal agents, even in cases where a "more educated" view would abandon such an approach in favor of a probabilistic (non-causal) analysis. While a statistician would likely accept some deterministic, and therefore causal, explanation for many chance events (e.g., flipping a coin), these explanations would be viewed as not particularly useful in some circumstances in that the causal explanation is overwhelming complex and practically unknowable. For this reason the "black-box" approach would be regarded as providing a more useful model for understanding. Furthermore, values associated with a prediction which is based on the explication of causes are apt to be a gauge of the person's confidence that the predicted outcome will occur. The value also seems to be interpreted as a measure of the strength of or degree to which causes are present.

In the next section the interrelationships among the three features of the outcome approach will be explored in an attempt to show that they may be regarded as interactive components of an informal approach to probability.

Coherence of the Outcome Approach

In the previous section the outcome approach was introduced and described as consisting of the single-trial, decision, and causal features. Excerpts were offered both in support of the validity of these features and to show how they are logically connected and thus can be considered core components of an informal approach to probability. However, the development of logical connections among the features is a necessary but not sufficient step in demonstrating their interrelatedness. In this section, two types of empirical data which could provide evidence of the interrelatedness of the features will be presented. First, the degree to which aspects of the features a) are consistent across problems and b) cluster together in subject protocols will be estimated by calculating relevant correlation coefficients. The first set of correlations provide an estimate of how reliable the features are. These correlations address the question, for example, of whether a subject who makes an evaluative (right/wrong) response in the Bone problem is more likely to make a similar response in the Weather problem. The second set of correlations address the question of whether the features occur together such that, for example, a subject who adopts a causal approach to determining a probability is more likely to make an evaluative response in the context of the occurrence of a predicted outcome.

The second source of empirical evidence will involve further qualitative analysis of the interviews. This analysis will focus on four protocols from the Weather problem in which the relations among the features and their incompatability with a frequency interpretation are especially apparent.

Correlations Within and Among Features

Within features. It was possible to determine the consistency of the causal feature across all three of the problems. However, the only other response-type that occurred in more than one problem was the evaluation of predictions which was observed in both the Weather and Bone problems.

Scores for the evaluative response in the Bone problem were generated by scoring 2 points if subjects made at least one statement to the effect that their prediction was right or wrong. In the case of the Weather problem, corresponding scores were generated by scoring a) 2 points if subjects said that the forecaster was wrong having predicted rain when, in fact, it did not rain, and b) 2 points if the 7-out-of-10 correct predictions were regarded as "pretty accurate" or 1 point if subjects were undecided between the "pretty accurate" and "perfect" evaluation. (See Table 3.)

Scores for the causal feature in the Bone problem ranged from 0 to 4, 1 point being scored for each category that was checked in Table 4. Corresponding scores for the Weather and Misfortune problems ranged from 0 to 2. For the Weather problem, 1 point was scored for each category that was checked in Table 4. For the Misfortune problem, a score of 0 indicated that only a chance explanation had been given or had been preferred to a non-chance explanation; 1 if both a chance and causal explanation were given but no preference was expressed for the former; and 2 points if only causal explanations were given.

Product-moment correlations were calculated and the significance of the values were evaluated using one-tailed t-tests. The correlation between the evaluative responses in the Bone and Weather problems was $r = .423$ which, however, was not significantly different from 0. Correlations among the causal features were: Bone-Weather $r = .634$, $p<.005$; Bone-Misfortune $r = .477$, $p<.05$; Weather-Misfortune $r = .460$, $p<.05$.

Among features. As has been mentioned, in the analysis of the interviews, the decision and single-trial features were not analyzed or coded separately since their co-occurrence was virtually perfect. Accordingly, the correlations which are examined in this section are those among the a) evaluative and b) yes/no predictions associated with the decision feature and c) the causal feature. Scores for the among-feature comparisons were obtained by combining the features scores of the individual problems. The evaluation-of-prediction scores ranged from 0 to 6, the scores for the Bone and Weather problems having been added together. For the causal feature, the scores on the Weather, Misfortune, and Bone problems were added together, giving a possible range of from 0 to 8. The yes/no prediction scores were based only on performance in the Weather problem. They ranged from 0 to 3, 1 point being scored for each category checked in Table 3.

Product-moment correlations were calculated and evaluated as in the within-feature analyses. The causal and evaluative responses were highly correlated ($r = .762$, $p<.0005$). The yes/no predictions and causal responses were also related ($r = .443$, $p<.05$). However, the yes/no predictions and evaluative reponses did not prove to be significantly related ($r = .275$).

Considering the restricted sample size and the fact that the feature scores were based on very few criteria, the correlational analyses provide evidence which suggests that, as a whole, the features as decribed are both consistent across problems and cluster together in subject protocols.

However, even strong correlational evidence that the features are consistent and cluster together would not necessarily demonstrate that they are core components of an informal approach to probability. It may be that they are common but unrelated notions held by those who are unfamiliar with formal probability theory, and their co-occurrence simply reflects this lack of formal instruction. What must be shown is that they not only cluster together, but that they are imbedded in a logical system which has an ontological status on par with a formal theory of probability. This demonstration is attempted in the following section.

Analysis of Conflicts

In this section the performance of four subjects on the Weather problem will be further examined. These subjects, as was previously mentioned, expressed conflicting ways of interpreting the question of the accuracy of a forecaster who had predicted 70% chance of rain for ten days, three days of which it did not rain. According to one perspective, the forecaster was perfectly accurate. However, from an alternative, and to some a more compelling viewpoint, perfect accuracy would have entailed rain on all ten days. At the heart of this conflict is the question of whether the forecaster is trying to formulate a) an accurate prediction of the relative frequency of rainy days, or b) a decision about whether or not it will, in fact, rain. The former objective is consistent with a formal approach to probability; the latter has been described as the primary objective of the outcome approach to probability. That subjects who entertain both of these perspectives experience conflict suggests that the two perspectives are of similar ontological status--that the outcome approach is an alternative system for dealing with uncertainty.

In the protocol of Subject 5 all three of the outcome approach features are explicitly mentioned. This subject had initially given a causal interpretation of the meaning of 70%. He evidenced some confusion when asked what he would conclude if it did not rain on a day for which 70% chance of rain had been forecast. At this point he constructed a very accurate frequentist account, drawing on the use of the frequency data in the Bone problem to help him think about the Weather problem:

> S5: In the same way that I used that [refers to 1000 data] to predict here [results of 10 trials]. They have years and years of back-log information--radar pictures--that they keep updating and they say this pattern has this...X%, X number of times it's become this....

By the time the accuracy probe was interjected, he had given both a statistical and a causal interpretation to the 70%. His first response to the probe was consistent with a frequency interpretation. But immediately he verbalized an opposing "intuition" according to which 70% was interpreted as meaning that it's going to rain. It is clear that he views these as incompatable interpretations, and he eventually reverts to the causal view:

> S5: He was absolutely on the dot, because it was a 70% chance of

rain and...it was accurate--seven days it was, three days it wasn't. That's seven out of ten--that's 70%. But generally when they say 70% chance of rain, that means--"70% of the time we've seen this, it's happened." I don't know, I just have this intuition that they mean more than--that the 70% means something other than just seven out of ten days it will rain. But I don't know what it means. I just know that if I saw 70% chance of rain, I'd wear a rain coat the next day, and I would probably be right.

I: Any idea of what, in addition, that's communicating, other than seven out of ten days it will rain?

S5: OK. Well it must have something to do with cloud formation or--there are a lot of things that are involved in weather. There's the barometric pressure, the atmosphere...the temperature....

In terms of the outcome approach, the logic of this subject's argument is as follows: Seventy percent means more than that it will rain 70% of the time--it means it will rain (yes/no predicting) and the prediction is probably right (evaluation of the prediction). Given that 70% can be interpreted as "rain," it must be derived from a measure of factors that produce (cause) rain rather than from frequency data. As was suggested earlier, this latter belief may be based on the notion that since frequency data do not cause rain, they cannot be used to predict rain.

Similar to Subject 5, Subject 6 had previously given a frequency interpretation of 70% chance of rain. When given the accuracy probe, however, her response was that the forecaster was "pretty good." Asked what would have happened if the forecaster had been perfectly accurate, she responded:

S6: Then it would have rained every day. I guess when they say 70% chance, they're only saying a chance. But they still predicted that there'd be rain just because there was that chance for it. You know, it wasn't like a 20% chance of rain, where they would have predicted that there would be no rain, even though there was that small chance.

I: So on the days it didn't rain, what should the weather forecaster have predicted?

S6: You know, a lower chance of rain. I can understand if they predicted a 60% chance and it still didn't rain, 'cause there's still 40% that it wouldn't rain. But they should--I would think it would be very difficult to exactly predict yes or no when they're only 60%. So he should have probably had something lower than a 70% chance though on the days that it didn't rain.

I: When you said that it's very difficult when they're predicting with 60%, what--?

S6: Yeah, if by their methods they've come to see that there's a 60% chance of rain, it's hard for them to say if there would be or not. Just 'cause they know that there is a 60% chance of rain, but they also know that there is a 40% chance of no rain. So it's hard to say it's going to rain. They can really just tell you the chances, and you have to predict for yourself.

This subject verbalizes the difference between a probabilistic proposition and the prediction of an outcome based on such a proposition. The distinction is problematic for intermediate values where the resulting prediction is ambiguous. For more extreme values like 70%, she regards the prediction as obvious, falling neatly within the "yes" region of the decision space. Her statement that "exact" predictions of yes or no cannot be made for intermediate values implies that unequivocal predictions can be made for more extreme values. Given this belief, one would be justified in saying that a forecaster had made a mistake if it did not rain on a day for which a 70% chance of rain had been assigned.

Subject 8, in response to the accuracy probe, concluded that the forecaster was "very accurate," then immediately began to wonder whether one could look beyond an individual prediction in assessing accuracy in this instance:

S8: I was just wondering whether--if you're taking a ten day span and, as you said, three of the days it didn't rain, if that can really relate to when he's looking at an individual day--that particular day. And I suppose it can if you're looking at a ten-day span with 70% chance of rain every day, with the same set-up.

When asked if the forecaster could have been any more accurate she apparently encoded the question as, "Should the forecaster have predicted higher than 70% chance?" She responded that giving higher proportions does not imply higher accuracy:

S8: No, I don't think so. No, 'cause when he says there's going to be 70% chance of rain, he really can't be more--I don't think he can be more predictive than that. 'Cause that's a proportion. If he said 50% chance of--that's, you know, not any more accurate than if he said 70% chance of rain.

I: Saying 50% chance of rain is no more

S8: Yeah.

I: accurate? In terms of what?

S8: I guess, as I said, he must have certain standards to go by when he picks a chance of rain, and--

In an attempt to get her to clarify her statements, she was further probed:

I: Say he predicted 50% chance of rain on those ten days. What would have happened if he were a really good forecaster? If you kept track, what would you expect in terms of the number of days it would rain and the number of days it wouldn't rain?

S8: Actually, you couldn't really expect anything. Because he is looking at an individual day, 50 and 50%. So let's say if it rained that day, then he had a--he was--I don't know, 'cause you're looking at individual days, really, so it could have rained all the time, it could have rained not one day out of ten days, and then it could have been 50/50, like five days it rained, five days it didn't rain, and he wouldn't be--and it would be the same, actually. It would come out the same, I guess, 'cause he is looking at individual days.

I: Tell me again what would come out the same? If over ten days it didn't rain at all--

S8: Yeah, and if it did rain. 'Cause he's looking at a particular day. And it's 50% chance rain, 50% not. So he wouldn't be more or less accurate in any of those situations, I don't think.

The above excerpt seems very confusing until the subject is viewed as interpreting the 50% in terms of the outcome approach. According to that view, 50%, as the mid point on a yes/no decision line, means that anything can happen--rain or no rain. Using this interpretation of 50% she confirms her conclusion that assessing accuracy over days is inappropriate since it should make no difference in the 50% situation what the results are for the ten days. If accuracy were judged according to the results of ten days, one would expect five days rain and five days no rain; but this interpretation of accuracy would contradict the use of 50% as an indicator of "anything-could-happen." This reasoning makes her more certain that individual days should be the unit of analysis, and accordingly, when asked to summarize what she believed, she responded:

S8: Well, he's looking at an individual day--particular day--and he's setting up percentages on one day. And you can't really extend that to an amount of time, I don't think.

Subject 11 experienced the same dilemma as Subject 6 of not knowing whether it is appropriate to assess accuracy by summing over days. He was not able to resolve the dilemma and moved back and forth between viewpoints in what he called a "logic swirl."

S11: Oh, it seems as though they were pretty much right on because in the whole of the ten days they were right on three of the occasions, which would be 30% it didn't rain, and that's what they kind of predicted, almost. Maybe.

I: Uh-huh. What are you thinking?

S11: Well, 'cause it's like--They didn't predict a ten-day span. They predicted each individual day. And so it seems as though when each day comes up, it's a whole in itself and it's not necessarily put together in a unit.

S11: ....It kind of gets back to the idea that they were pretty much right getting seven out of ten right. Maybe it's not fair to judge it that way. Maybe you should just judge each one. But I guess you can add them together because they're all the same. They're all like 70%....

I: ....Would they have been more accurate had it rained on all ten of the days? Would that be more impressive to you?

S11: Well, that's weird 'cause it almost seems that--it almost seems, I don't know.

I: What are you thinking?

S11: Well, I'm thinking that if they go over, it's almost like they're going over and they're wrong the other way. It's raining more than they really predicted. But it's not like they predicted that it would rain on 70% of the days--of the next ten days. It's like they predicted rain for each day. And if they were 70% sure...that it was going to rain on each day...if it rained on more of those days, then as you increase the number above seven, then they would be less accurate.

I: So you're sort of in a dilemma there?

S11: Yeah. I don't know. It depends.

In this chapter I have constructed an informal approach to probability according to which subjects' responses to the problems could be viewed as reasonable. The informal approach has been described as an outcome approach since the primary objective is to predict what will or did occur on a particular trial. Arriving at a prediction often involves an analysis of the presence of factors which would tend to cause individual outcome alternatives. Numeric values that may be associated with a prediction are measures of the confidence that the predicted outcome will occur, as well as measures of the extent to which causal agents predispose that outcome. When probabilities are given to the outcome-oriented individual, they are interpreted consistent with the two characteristics mentioned above and are recoded into a yes/no or I-don't-know decision according to their distance from the corresponding values, 100, 0, and 50%.

Testing the Outcome Approach: Some Predictions

In the above discussion I have argued that many people, and to varying degrees, have an outcome approach to situations involving uncertain outcomes. The test of the validity of this hypothesis involved various tests of "internal coherence." First, various indicators of the outcome approach were found to be highly correlated both across subjects and problems. Second, analyses of interview protocols indicated that the characteristic features that were highly correlated plausibly came from a common source. However, since the outcome approach was inferred in large part from the same data set, a stronger test of the validity of the approach seemed desirable. To this end, I arranged for the same subjects to participate in a second set of interviews during the following semester. For these follow-up interviews, four new problems were used to test various predictions derived from the outcome approach. In this section, the new problems will be briefly introduced. The problems, along with the predicted responses, are described below. The predictions state how outcome-oriented individuals will perform relative to frequentist-oriented individuals.

Cab Problem

The Cab problem was used to determine the correspondence between the deterministic approach (cf. Well et al., 1981) and two features of the outcome approach. In the Cab problem, subjects are asked to estimate the probability that a cab that was involved in a hit-and-run accident was blue rather than green. They are given the relative number of blue and green cabs in the city as well as information regarding the accuracy of a witness who identified the cab as blue. According to the deterministic approach, some subjects feel that the task in the Cab problem is not to estimate the probability that the cab was blue, but to determine whether or not it was blue. This corresponds to the evaluative component of the decision feature and is characterized in the Cab problem by subjects making statements to the effect, "I think it was a blue cab," or, "I think the witness was right." These same subjects, according to Well et al., more frequently asked if a number was required, and when they finally gave an answer, it tended to be based on a "loose" or qualitative interpretation of the source of evidence they thought relevant. Thus, subjects often stated that the witness was "pretty good," giving a value near but not identical to the 80% when asked for the probability that the cab was blue. As argued above in the description of the outcome approach, this response is evidence of the tendency to predict in yes/no fashion.

Prediction 1. Outcome-oriented individuals will more frequently give those responses described by Well et al. as deterministic in answering the Cab problem.

Bone-2 Problem

Relatively few subjects (three) were judged in the Bone problem to have made predictions about rolling the bone from its physical features. This may have been because, not having expected this response, there were no probes designed to elicit it. It may also be that (as expressed by Subject 9) many felt unable to make an educated analysis of the physical features, though they believe that it is through such an analysis that probabilities would ultimately be determined. To explore this issue, in the Bone-2 problem subjects are asked to predict the outcome of another ten rolls of the bone, and are given the choice between making predictions on the basis of frequency data or on the basis of information related to the physical properties of the bone. They are also asked how a statistician would determine the probabilities associated with each surface of the bone.

Prediction 2. Outcome-oriented individuals will more frequently choose information related to the physical properties of the bone in making predictions and state that a statistican would use such properties in determining probabilities.

Painted-die Problem

In the Painted-die problem, a population of five black stones and one white stone together with a sampling-with-replacement procedure is described. Subjects are asked to predict whether in six trials it is more likely to sample six black stones or five black and one white. According to the outcome approach, the primary task in situations of uncertainty is to predict the outcome of individual trials. When asked to predict the overall outcome of a series of trials, the outcome-oriented individual will not use the fact that there are five out of six black stones to form an expectation consistent with probability theory that they will sample, on the average, one white for every five black stones. Rather than interpreting the 5/6 figure as a probability which can be used as a basis for prediction, the outcome-oriented individual will interpret it in a more qualitative way as overwhelming evidence for black on each trial. A belief that in repeated samples they would observe more than five black for every white is evidence of this qualitative interpretation.

Prediction 3. Outcome-oriented individuals will more frequently state that in a sample of six stones, it is more likely to get six black stones than five black stones, and their reasoning will be that there are so few white. When asked how many white stones they would expect in 60 trials, they will more frequently give a number less than 10. Also, when actually sampling from the above population, they will give "probability values" greater than the percent corresponding to 5/6.

Modelling Problem

Subjects are asked in the Modelling problem whether some number of labeled stones could be put into a container and sampled in a way that would generate results comparable to rolling the bone. According to the outcome approach, frequency data are caused by features of the bone. If the causal features of two processes are different, as they appear to be in the case of rolling a bone verses drawing stones from a container, they will believe that it will be reflected in differences in the frequency data generated by the two processes.

Prediction 4. The outcome-oriented individual will more frequently state that it is not possible to model the rolling of the bone by randomly drawing labeled stones from a container.

CHAPTER IV

METHOD: INTERVIEW 2

Subjects

Twelve of the original sixteen subjects returned to participate in the follow-up interviews. The other four could not be located. Subjects were offered a choice of either extra course credit or $10.00 for participating. Approximately five months had elapsed between the first and second set of interviews. Additional math or statistics instruction that subjects had received during this interval is noted in Appendix A.

Procedure

Setting

Initial instructions to subjects were nearly identical to those outlined in Chapter II. Subjects were told that they would be given several new problems that were similar to the last set in that they involved predicting outcomes in situations of uncertainty. They were reminded to "think aloud" and that they could use the pen and paper for any figuring they might want to do.

Problems

Four problems were employed to test the predictions as described in the previous chapter. The Cab problem had been used in previous research as earlier noted. The remaining three problems were developed and then standardized in 14 pilot interviews. All four problems are presented below in the order they appeared in the interview.

Cab problem. (Subjects were asked to read the Cab problem aloud.) "A cab was involved in a hit-and-run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

(i) 85% of the cabs in the city are Green and 15% are Blue.

(ii) A witness identified the cab as a Blue cab. The court tested his ability to identify cabs under the appropriate visibility conditions. When presented with a sample of cabs (half of which were Blue and half of which were Green) the witness made correct indentifications in 80% of the cases and erred in 20% of the cases.

What is the probability that the cab involved in the accident was Blue rather than Green?" (After subjects give a numerical response they are asked:) How did you arrive at that number?

Suppose the information in (i) were reversed such that 85% of the cabs in the city were Blue and 15% were Green. The witness, as before, identified it as Blue and was 80% correct in the test situation. In that case, what would you say the probability was that the cab involved in the accident was Blue?

Bone-2 problem. Last time you were asked which side of this bone you thought would most likely land upright. Do you remember which side you concluded? (The bone is held far enough away so that the labels can not be read.) I'm going to ask you the same question again. And to give you something to base your answer on, I'll offer you any one of the following pieces of information. (Subject is shown the list as the interviewer reads the items.)

1-A measure of surface area of each side.

2-The results of 100 rolls made by 16 people.
3-The results I got in 1000 rolls.
4-A drawing of the bone showing the center of gravity.
5-The bone to look at.
6-The results on your last 10 rolls.

Which one would you like? Why did you chose that? If you could have a second piece of information, which would you choose? Why did you choose that? (Subjects are given both choices unless item 4 has been picked. In that case, they are told that the drawing is not available, and to pick another item. The estimate of surface area is in square inches: A=.028, B=.078, C=.065, D=.169, E=.018, F=.031. The results of 100 rolls were: A=7, B=32, C=21, D=35, E=5, F=0.) If you rolled the bone, which side do you think would most likely land upright? What is the probability that x will land upright? (Subjects are asked to predict what they would get in ten trials, then the trials are conducted as in the first interview.)

I have here a six-sided die. Suppose I told you that there was a possibility that it was loaded--that it had been altered so that one side was slightly more likely than the others to come up. Could you determine whether or not it was loaded? How? Would rolling it help you determine whether it was loaded? Suppose you rolled it 24 times and got the following results: (Subject is shown the results as the interviewer reads them.) 1=5; 2=2; 3=8; 4=2; 5=4; 6=3. What would you conclude?

In fact, the die is not loaded. Suppose I painted five of the surfaces black and the other one white. If I rolled the painted die six times, would I be more likely to get six blacks or five blacks and one white? If I rolled it 60 times, how many times would you expect the white surface to come up? (This probe was originally worded, "On the average, how many times would you have to draw from the cup until you got a white?" After the third interview, it was changed to the present form which was easier for subjects to understand.)

Obviously, I haven't painted the die. But I do have five black stones and one white one. (The stones were identically shaped pieces from a board game which subjects had used in a previous problem in the first set of interviews.) Suppose I put these in this cup and shook it up really well. Then I reached in without looking and drew one out, wrote down the color, replaced it, shook it up again and kept drawing like that. (This is demonstrated as it is explained.) Would that be the same as rolling the painted die? If I rolled the die several times and recorded what I got, and I drew stones and recorded those results, could you tell from looking at the results, which I got from rolling the die and which from drawing stones? I'm going to draw six stones from the cup, but first ask you to predict what I'll get? (Stones are sampled, and before shown the results of each trial, the subject is asked to say which color has been drawn, and also the probability that it is that color.)

Modelling problem. You agreed that we could create a model of the painted die by drawing stones from a certain cup--that that would give comparable results. Would there be a similar way that we could make a model of the bone so that instead of rolling the bone we could pick something out of a container and get the same kind of results?

(Subjects are given the following probes successively until they agree upon a model or reached the end of the list:)

1-How about if we put six stones which have been labeled A through F in this cup and sampled from it as we did before?

2-Is there some container that I could fill with some number of lettered stones that would give results similar to rolling the bone?

3-Suppose we took the bone to a statistician and, however it is done, the following probabilities were calculated for each side: (Subjects are shown the list as the interviewer reads it.) A was 5 out of 100, or 5%; B was 29 out of 100, 29%; C, 24; D, 37; E, 5; and F, 0. So we took a big can and first put five of these stones which have been labeled A inside. (A large can and six small containers filled with labeled stones is produced.) Then we took 29 Bs, 24 Cs, 37 Ds, and 5 Es, and put them in the container also. Then we shook it up and sampled from it as before. Do you think that would give results comparable to rolling the bone?

4-Suppose we rolled the bone and, say, we got a B. We took a stone labeled B and put it in the container. Then we rolled the bone again, and similarly, whatever we got, we put the appropriately labeled stone in the container, and we did that over and over. Would we reach a point when it would make no difference if we rolled the bone or drew from the container we had filled?

(When, and if, subjects agree upon a model of the bone, they are asked the following questions:) If I rolled the bone 100 times and kept track of what I got, then I drew 100 times from this can filled with the labeled stones, and I showed you the results from both, could you tell from looking at the results, which I got from rolling the bone and which from drawing from the container? In those 100 trials with the bone and the container, do you think with one of those I'd be more likely than with the other to get no Es? Do you think I'd be more likely with one of those to get more Ds in 100 trials than with the other?

Presentation

The problems were presented orally by the interviewer, and in the order as specified above. The interview format was the same as in the first set. A complete transcript of the interview with Subject 11 is included in Appendix C.

The interviews required approximately 40 minutes. Following the interview, questions about any of the problems in either session were answered as were questions concerning the purposes and preliminary findings of the research.

CHAPTER V

ANALYSIS AND DISCUSSION: INTERVIEW 2

Four predictions were made to test the validity of the outcome approach as described in Chapter III. For the purpose of testing these predictions, scores based on performance in the second set of interviews were correlated with an outcome score which summarized subjects' performance on the first set of interviews. Outcome scores were generated for each subject by summing their individual feature scores as described in Chapter III. Outcome scores had a possible range of from 0 to 17, higher scores being indicative of an outcome orientation. For the 12 subjects who were interviewed on the second occasion, outcome scores ranged from 0 to 15, with a mean of 6.83 and standard deviation of 4.39.

Scores related to each of the four predictions were generated by having two raters independently code each interview. The raters included myself and a graduate student who was blind both with respect to the nature of the first set of interviews and to the hypotheses being tested.

In this chapter I will consider each of the predictions in turn, describe the coding and scoring procedures that were used, report correlations between performance on the two sets of interviews, and present selected excerpts from the interviews that pertain to the predictions.

Prediction 1

The Cab problem was included in the second set of interviews both to replicate the findings of Well et al. (1981), and to determine the extent to which responses labeled in their report as "deterministic" correlate with the outcome-oriented responses on the first set of interviews.

Performance of subjects on the Cab problem is summarized in Table 5, Appendix A. In accord with the findings of previous research, the modal response to the Cab problem was 80%. Subjects who gave 80% as the answer generally expressed the belief that the information concerning the base rates of green and blue cabs was irrelevant and should therefore not figure into the answer. Two subjects (6 and 8) argued that both the witness and base-rate information were important to consider and gave estimates of the probability that were intermediate between the 15% and 80% (.60 in both cases). These results are in basic agreement with those reported by Well et al.

It was predicted that outcome-oriented subjects, as definded by higher outcome scores on the first set of interviews, would respond to the Cab problem in a way that was characterized by Well et al. as deterministic. Specifically, it was predicted that outcome-oriented subjects would be more likely to:

1) ask whether a number was required in answering the question of the probability that it was a blue cab;

2) encode the question, "What is the probability..." into the

question, "What color was the cab?" (this recoding being indicated by responses to the effect, "I think the cab was blue"); and

3) base a numeric answer on a "loose" or qualitative interpretation of the evidence they thought relevant.

Performance of subjects with respect to these three predictions is summarized in Table 5. Interrater reliability for coding these three categories was estimated by correlating the set of ratings of the two coders, with r = .759. The scoring rule applied was that both coders had to agree that a particular statement had been made in order for it to be counted. Scores on the Cab problem were obtained by assigning 1 point for each of the three categories checked, and summing across categories for each subject. Thus the scores could range from 0 to 3. These scores were correlated with the outcome scores (as indicated in Table 5) with $r = .612$, $p<.025$ (one-tailed). Considering the moderate level of interrater reliability, the magnitude of the correlation coefficient provides strong support for the argument that the deterministic responses on the Cab problem are related to the features of the outcome approach. That the predicted responses to the Cab problem are characteristic of the decision and single-trial features of the outcome approach is also supported by the content of the interview excerpts which are provided below.

Given that the goal of the outcome approach is to determine what will or did occur, the question of probability is translated into the question, "What happened?" as indicated in the following response:

S1: So you want to know if I think that's right--if it was blue.

> Well, I would say that it would be blue rather than green--just the fact that this really isn't important--the 85% are green, 15% are blue. I mean there are still a substantial amount of blue cabs out there. But the fact that the guy said--well, the court said that in 80% of the cases you identified the right color. And the guy said he saw blue. He doesn't say, "I think I saw blue." He says, "I saw blue." So I would go with blue.

Given that the Cab problem asks specifically for the probability that the cab was blue, subjects' query of whether a number is required is indicative of the decision feature. When subjects asked if a number was wanted, I hesitated in order to allow them to clarify the question, and then if they did not continue, I asked them what the alternative was to giving a number.

> S11: Let's see. Am I looking for a number as opposed to like--Am I looking to say "it's 80% probability that it was a blue rather than green? Is that what I'm--
>
> I: What's the other option? How else would you prefer to give that?
>
> S11: "Sure, it could have been a blue cab." No--just that it would have been a strong--it was more likely as opposed to less likely. Kind of like this fit in. More positive as opposed to a definite number positive.

Central to the goal of specifying what happened is the single-trial feature, according to which questions of uncertainty are viewed as pertaining to a particular event as opposed to a set of events. Subject 5 justified ignoring the base-rate information on the grounds that at issue was the occurrence of a particular event, and that information regarding a class of events was irrelevant:

> S5: It really doesn't matter how many cabs there are in the

> city. What you're thinking about is this one particular cab, whether it was blue or green. And since the guy was usually right--he's probably right.

As indicated above, the witness identification can be seen as applying to the individual event (the color of the errant cab) in a way the base-rate information cannot. Using the base rates would seem to require regarding this particular accident as one of a set of accidents involving the two cab companies. To the outcome-oriented individual, this is not relevant to the question. What matters is this particular accident. (On this basis it could be predicted that outcome-oriented individuals would be less likely to use base rates than individuals employing a frequency interpretation. Not enough data are yet available to test this prediction.) It is evident in the above and following excerpts that the witness identification is not viewed as one of a class of similar identifications. Rather, the test situation allows the outcome-oriented individual to assign the attribute "pretty reliable" to the witness and thus to the witnesses' identification of the color of the errant cab. It is in the process of making this attribute that subjects "let go" of the specific meaning of the 80% and then given a confidence value for their belief that the cab was blue which is only loosely based on the 80% estimate of the witness's accuracy.

> S8: And since his visibility was pretty clear, and just on that--I'm not even taking these numbers so much as just, you know, conceptualizing it. Since he saw it was blue and there's more of a chance that he's right as seeing it as blue, that he saw it correctly. So I'll say that.

S3: 80% just because he had--his percentage correct before was 80%, so it makes sense that he, probably--chance 80% that he got it right this time.

I: OK.

S3: Maybe better.

I: Maybe better?

S3: Yeah. Just--I--

I: Can you explain why you think it might be better than that?

S3: Well, because more than not he got them right when they tested him before. So that's why it would be possible that he'd be more than 80%. I don't know if that makes sense.

I: About how much more, do you think? Maybe you could just give it a range.

S3: Oh, 5 to 10%.

I: So maybe 85 to 90%?

S3: Yeah.

S12: If someone actually saw that it was a blue cab, then I'd assume it was a blue cab.

I: And what would be the probability that it was blue?

S12: I would say about 70%

I: And how are you getting the 70%, approximately?

S12: Well, allowing for a 10% error in the statistics of 80% of the cases were correct and 20% were error. So I'm giving this a little more--you know.

I: And why do you feel you want to do that--give a little bit more--room for error?

S12: I'm just a nice person.

S13: Yeah--that he did guess, more than he didn't, the right

colors. So I'd go with the blue. I'd say that it was a blue one.
[The coders disagreed as to whether this was an instance of a qualitative response.]

I: And how about just an estimate of what the probability would be, or a guess.

S13: I want to say just 80....

I: Is that 80 based on this [points to 80% witness accuracy]?

S13: No. I'm just trying to find--I'm just trying to think of something that's closer to 100--like over to more of a chance that it happened.

Prediction 2

The second prediction was that outcome-oriented subjects would prefer the physical features of the bone to frequency data in arriving at a judgment of which side of the bone was most likely to land upright. Since this tendency was not anticipated before the first interview was conducted, there was at that time little probing included to explore this behavior. Consequently, only a few subjects in the first interview were judged to have shown such a preference. The Bone-2 problem was designed to test the extent to which subjects prefer a physical to a statistical analysis of the bone in making predictions of its tendency to land on its various sides. Subjects were required to a) choose from among six alternatives the particular information they would prefer in anwering the Bone-2 problem, b) to predict the results of ten rolls of the bone (the information they used to derive their predictions being noted), and c) to say how they thought a statistician would determine the probabilities of each side of the bone landing upright.

Subjects' performance on the Bone-2 problem is summarized in Table 6 (Appendix A). Scores on the problem were obtained by scoring a) 2 points if subjects' first choice was not a frequency and 1 point if the second choice was not a frequency; b) 2 points if only physical properties were used in predicting ten trials, 1 point if used in conjuction with frequencies; c) 2 points if only physical properties were mentioned in the context of a statistician's method, 1 point if mentioned along with frequencies. Scores had a potential range of from 0 to 7. The interrater reliability on this problem was 100%. The correlation between scores on the Bone-2 problem and the outcome scores provide evidence of a very strong relationship, with $r = .788$, $p<.005$ (one-tailed).

Two possible reasons for preferring a physical to a statistical analysis were suggested in Chapter III. One possibility was that the physical features of the bone might be viewed as a more stable source of evidence when compared to frequency data which can fluctuate from sample to sample. This seemed to be the rationale given by Subject 12 for basing predictions on an inspection of the bone. Asked why she thought the data from 100 rolls was unimportant, she replied,

> S12: Well, because what they did may not be--it's sort of chance, you know, that happened. If the same 16 people did the same 100 rolls, it would probably be different the second time. It just doesn't seem a very specific kind of statistic.
>
> I: And why do you think it would be different?

S12: Things change. I don't think anything duplicates itself exactly the second time.

I: How about the results I got in 1000 rolls?

S12: Yeah, that too is kind of iffy. If you did the same thing over again, plus a second 1000 rolls--I mean you could go on for 2000 rolls or whatever, and I don't know if it really would tell you much. Then again, I could be wrong.

The second hypothesis as to why a physical analysis might be preferred is that physical properties are viewed as causal agents of what one wants to predict while frequency information is not. The first set of interviews provided no compelling evidence that this is the case; nor did the second. One subject did feel that the physical properties were "real evidence" in contrast to frequency data. Asked to explain how she decided that D was the most likely side she responded,

S3: Well, just 'cause it's flatter on the underside, so it's more likely to land on that side than it would on any other place.

I: Are you using this information at all [the results on her last ten rolls]?

S3: Maybe a little, yeah. I suppose. Well, I looked first and thought that was reasonable. So--

Asked how a statistican would determine the probability, she first mentioned surface area. Asked if they would use anything else, she said,

S3: Well, they would probably make rolls themselves and see how it comes up. But I don't know if they would use that for real evidence or whatever.

I: You feel like the results of what you got isn't real evidence?

S3: Well, yeah. It has some. But there must be some, you know, like measuring the side, and that must be a little more precise than my rolls.

Why she regards the properties of the bone as a more valid type of evidence is not clear. Her last statement above, however, suggests that it may be because they are easier for her to think of as being measured precisely.

The other three subjects who considered features of the bone to be important in a determination of probabilites felt that they should be used in conjunction with rather than to the exclusion of frequency data. Given the six alternatives, Subject 13 chose the frequency data, using it to arrive at the decision that D was most likely, with 70% probability.

S13: I'm just saying 70 because--Well, you get 35 for D, 32 for B, 21 for C [from n=100 data] and, I don't know. Like B would come in second, right? In my thinking. C would come in third and then A and E. But 70 sort of gives you over 50/50 chance, and it sort of gives you leeway to--maybe 75--sort of gives you leeway that you could still--it won't come up, and that the 32--either B or C--

When asked to predict the results of ten trials, she quickly wrote zero for both E and F, then after a pause asked,

S13: I can't look at it [the bone] can I?

I: Would you like to do that too?

S13: Yeah. I forgot which one--where they are.

After she had made her predictions, she was asked how she arrived at the particular values.

S13: ...D and B, because--I gave those three each because I just think that they'll come up most often 'cause of the area that they take up on it. Plus, too, that they have--they came up most often in these ones. And the A and E, I think they have an equal chance.

S13: ....C came up more often, though, on these. [Inspects bone] I don't see why. Both look--maybe I should--No, I'll leave it like that.

She later, in fact, revised her prediction so that C was expected more frequently than A in accordance with the frequency data. She appeared to use the two sources together, wanting to make sense of the frequency data by inspecting the surfaces of the bone.

Subjects 5 and 11 used the two sources of information in a similar way, analyzing the physical features in an attempt to explain the frequency data and, in cases, to arbitrate close decisions.

S5: ...I'd take number three [n=1000 data], and I'd look at each surface of the bone that had come up and compare it to the number of times it had gotten up and see why it had so I could decide whether or not the results were accurate, according to the shape of the bone.

In predicting ten rolls Subject 5 inspected the bone carefully to decide how he would alot rolls to B and C since, according to him, they were so close in the frequency of past occurrence. His explanation of how a statistican would estimate probabilities was consistent with the approach he had employed:

S5: ...a statistican would count a great deal of weight to the center of gravity and how it related, and, taking your results [n=1000 data], would come up with a bunch of statistics that would probably reflect fairly accurately your results, with perhaps some modification according to what he thought the structure of the bone gave out.

Subject 11 used only frequency data to make predictions about the bone, but felt that a statistican would, in a "joint effort," supplement these with an analysis of physical properties:

S11: 'Cause you'd roll the bone and get a rough idea of the probabilities, whatever they are--yeah, probabilites--and take it to have it analyzed to figure out if, structurally, you can understand why these--you know. You assign these particular values to each face, and then through comparing both, just--

I: But I might want to modify what I had got rolling it?

S11: Yeah. It's just kind of like added significance, or not significance--added sureness, or whatever--belief in your percentages.

In summary, the tendency to view physical properties of the bone as important in the determination of probabilites of the various landing orientations is highly related to measures of the outcome approach. Physical properties appear to be regarded as information at least on a par with frequency data in making predictions.

That the first two predictions were verified is not surprising. Basically what has been established is that subjects will respond consistently vis-a-vis the outcome approach on problems of the same type. The last two predictions are meant to serve a different purpose. They involve using the outcome approach to anticipate specific responses that had not yet been observed. Because of this, they provide more compelling evidence of the validity of the outcome approach.

Prediction 3

In the Painted-die problem subjects were first presented a die and then six stones, both of which consisted of five elementary outcomes of one type (black) and one of another (white). They were asked to predict whether in six trials they would be more likely to observe five blacks and one white, or six blacks. The answer according to formal probability theory is that the former is more likely, the probability of exactly five blacks being .402, the probability of six blacks being .335.

Most people, when asked, will respond that the probability of white being drawn is one out of six. But it is not clear what is meant by such an answer other than that there is only one white out of six sides. If people viewed "one out of six" consistent with formal probability theory they would expect to get, on the average, one white in six trials. That is the modal outcome. Even failing this line of reasoning, one would predict on the basis of the representativeness hueristic (cf. Kahneman & Tversky, 1972) that people would believe five blacks to be the more likely outcome. Since five black and one white looks more like, and in this case is identical to, the population distribution, it should be regarded as more likely than a sample which does not as closely resemble the population.

It was predicted that outcome-oriented subjects, however, would regard six blacks as the more likely outcome. In the outcome approach the primary unit of analysis is the individual trial. Given a probability value, one way to arrive at a prediction of a trial is to decide which yes/no or I-don't-know decision point the probability value is closest to. Similarily, rather than viewing the 5/6 as a value that relates to the expected relative frequency of blacks in randomly drawn samples, it is given a qualitative interpretation of the approximate form "the next trial will result in a black." When one is therefore asked to predict six trials, rather than using the 5/6 to form an expectation for the set of six trials, they will arrive at a prediction by summing over their expectations for each of the six trials. Since this expectation is more qualitative than quantitative in nature, it is expected that outcome-oriented subjects will more frequently say that six blacks are more likely since there are so many of them, and that they will also believe that the ratio of blacks to white over larger series of trials will remain above the normative value of five to one.

Interrater reliability for coding the Painted-die problem was 100%. Table 7 summarizes subjects' responses. Seven subjects felt that six blacks was more likely in the case of the die, six subjects in the case of the stones. (Subject 13 regarded the two differently and changed her prediction to five blacks in the stone problem.) Scores for the Painted-die problem were obtained by giving 1 point if six blacks were expected in either the die or stone version; 1 point if fewer than 10 white were expected in 60 trials or if more than six trials, on the average, were required to roll one white; and 1 point if the probability of a black on the first trial was estimated to be above 5/6 or 84%. The scores obtained in this manner were correlated with the outcome scores from the first set of interviews, with $r = .573$, $p<.05$.

Excerpts from the interviews suggest that, as predicted, subjects solved the problem by imagining a single trial for which the probability of black is overwhelming, and then extending this over trials to arrive at the conclusion that six blacks was the more likely outcome.

> S7: Well, I think it's--the white's there, but--I'm not exactly sure what trying to say. Just because the odds are always the same. There's only one of them in there. So even though it's six rolls and there's six things in there, there's only one or the other that's going to come up each time. And that--chances are better than--five to one, one of the five blacks is going to come up.

The responses of Subject 7 were characteristic of a frequency interpretation in the first set of interviews, and though he knows that the probability is 5/6 for a black, he still feels that six blacks are more likely and that ten or less whites would occur in 60 trials. Similar reasoning is demonstrated by Subjects 12 and 13.

> S13: I'd say the black have more chance of coming up. Black side.
>
> S13: ....Just, the odds don't seem to be favorable for one white.
>
> S12: I would probably just say you'd get six blacks just because there are more black sides. The probability of the white coming up would lessen, I think.
>
> I: Would lessen? What do you mean "would lessen"?
>
> S12: Would diminish.
>
> I: Compared to what?
>
> S12: Compared to the black sides coming up.

The first response of Subject 15 suggests that she is chosing six black because she doesn't think things are likely to turn out perfect:

S15: Because to get five black sides and one white side in six throws is 100%--is just perfect. And I think it's more likely that you'd get the six black sides.

It's clear from the rest of the protocol, however, that she has other reasons for her belief:

S15: Because it's a higher probability of getting a black side because there are more black sides and so there's more probability that when you roll it, you're going to get a black side instead of that one white side.

Subject 3 combined the "more blacks" rationale with the reasoning that the sampling with replacement procedure does not guarantee white.

S3: Probably more likely to get all black just cause--I don't know what percentage, but most of the die is black, so it's going to come up on that side. 'Cause you're not going to roll it on a different side each time you roll it, so that it's bound to come up one of those six rolls. So it probably would be black on all of them.

Subject 5 believed that rolling six dice at once would result in five blacks, but that rolling the same die six times would result in six blacks.

S5: Well, each roll is a separate entity. You roll it, and a side will come out. Then you roll it, and a side will come out. You don't roll all six at one time. So likelihood is that each time it comes out, the side that has the dominate color, which is black, is the color that'll come out.

He finally rejected this reasoning, favoring five blacks in both cases. His intial response, however, provides a good example of what is being regarded as the outcome approach to this problem--that of imagining the results of one trial as almost certainly being black, and by extending this qualitative judgment, concluding that six blacks are more likely over six trials. It is especially significant that what got this subject thinking differently about it was to imagine all six trials occurring at once. (A similar belief in a difference between flipping one coin repeatedly or several at once was defended by the 18th century mathematician, D'Alembert. For an interesting account of this and other of D'Alembert's unconventional beliefs about probability, see Todhunter, 1949.)

Prediction 4

The Modelling problem was designed to test some implications of the causal feature of the outcome approach. According to the outcome approach, frequency data are not considered to be as reliable a source in predicting outcomes as are phenomena that are causally related to the outcome. This being the case, it was predicted that outcome-oriented individuals would hold that if the causal features of a set-up were altered, outcome frequencies for that set-up would change accordingly. In the Modelling problem subjects were asked if it would be possible to construct a model of the bone that would produce results which could not be distinguished from results obtained from rolling the bone. They were introduced to the modelling concept in the Painted-die problem where it was suggested that randomly sampling with replacement from six identically shaped stones would be the same as rolling a fair die. It was assumed that most subjects would accept this comparison since the most obvious physical feature--the symmetry of the six sides--was maintained. With an urn model of the bone, however, the important physical aspects of the bone--its irregularly-shaped sides and unequal distribution of weight--are transformed into unequal numbers of objects which are identical in weight and shape. It was predicted that outcome-oriented subjects, focusing on this difference, would expect that the data obtained from conducting trials on the two set-ups would be distinguishable in some way.

The results of the Modelling problem are summarized in Table 8. Subjects' performance on the modelling component of the Painted-die problem is included as well. Scores for the Modelling problem were generated by assigning 1 point for each of the following four categories: a) if the urn model was not accepted in the case of the die; b) if a model of the bone could not be generated by the subject; c) if the can filled with numbers of stones suggested by the statistican's estimates was not accepted as a model of the bone, but the trial-by-trial method of filling the can was accepted, d) if it was believed that no model of the bone could be created. Scores on this problem could thus range from 0 to 4. Interrater reliability for coding in these four categories was determined by correlating the two sets of ratings, with $r = .93$. The correlation between these scores and the outcome scores was $r = .610$, $p < .025$.

The reasons given by subjects for rejecting the urn models are congruent with the hypothesis that they would do so because, in their analysis, important causal features could not be duplicated in the urn models.

Subjects 3 and 13 felt the urn model inappropriate in the Painted-die problem. They did not express concern over the corresponding features of the die and stone-filled urn per se, but over the differing sampling procedures in the two cases.

S3: I think maybe the white side of the die would come up more, just 'cause you don't have any control over that [makes an imaginary roll of the die]--Well, not that you do with the pieces, but--I don't know why. That's just what I think.

S3: ....because you're putting your hand in there and taking out. I just, I don't know why, but I don't think you'd pick the white one as often as the white side of the die.

S13: I just think grabbing something out--if you're grabbing it out, I think it would be more probable of being white. I don't know exactly why I'm thinking that way, but with this [die] I just [rolls die]--I don't know, tossing something just seems less of a chance, but picking something out seems more of a chance. You'd think it would be the other way around, though. But, I don't know....

In the following excerpts, subjects explain why an urn model is inappropriate in the case of the bone. The fact that the bone has six sides, uneven surfaces, and is rolled rather than picked out are all mentioned as important differences.

S3: Probably be more likely to get no Es with the container full of 100 pieces. Just--well there is a slighter chance that it would come up, and there's six sides. So that's why I think it's more likely to come up on the bone.

I: Because?

S3: Because there's only six sides....

S6: Probably it would be more likely to get no Es from the bone, 'cause the bone has to stand like that, and it would be easier just sitting in there--they don't have to--it's not like there's anything to do with the way it can stand and stuff like that.

S6: ...[D] might be more likely from the bone. I don't really think you can say, but it just might be just because the Ds are all mixed up in the can, whereas in the bone, that's the easiest side for it to land on. That's the most--that's the way it stands easiest, so you might get it more times in a row in the bone.

S7: ...you could easily pick 100 of them out without hitting an E. You'd have more trouble tossing the bone so you didn't come up with an E.

I: And why is that again?

S7: ...It just seems like because you're picking them out you could just miss one of the Es.

S11: ...there's so many in this group as opposed to the bone. I don't know, something about the bone and then seeing many, many, many, and many of these guys [Ds] coming up.

S11: ....[Looking at the bone and at the urn model] What I'm trying to decide is whether or not just numbers is a good way to represent the chances, or the probability, that these guys will come up in the natural situation. And, on one side it seems as though that they wouldn't because--Well, I don't know. Maybe they--I think it actually would because it just represents the chance of them coming up. And sure, there's a lot of these guys here, but there's--in the bone situation there's just a real high chance--or just through the natural rolling--that D and C and B will come up most often.

S15: ...these stones and the die are uniform, and each side is the same--it's the same surface. And this [bone] is all different. So this will affect--the shape of the side will affect the way it's going to roll. Like it would be harder for it to stand up on E like that. So you'd have to replicate the little indents and stuff like--So you couldn't make a--you couldn't turn it into six stones or something like that.

The persistance demonstrated by subjects in insisting that the bone could not be modelled was particularly impressive. The interview probes were designed to give subjects several opportunities to accept a model: They were given one alternative after another. The independent coder, not knowing the intention in this probing, discretely noted in two instances that the subjects had been strongly led to accept a model. The other subjects were as strongly "led" but insisted repeatedly that the model suggested would not be comparable to rolling the bone. Attending to the physical features as opposed to the resultant frequency data of a chance set-up appears to be a deeply ingrained orientation. Speculations about the origin and function of this tendency as well as of the other two components of the outcome approach will be offered in the next and final chapter.

CHAPTER VI

GENERAL DISCUSSION

This study was undertaken with the goal of inferring an informal approach to probability that would explain, among other things, why subject responses to problems involving uncertainty deviate from those prescribed by formal theories. On the basis of an initial set of interviews such an informal approach was hypothesized and described as outcome-oriented. In a second set of interviews, the outcome approach was used to successfully predict the performance of subjects on a different set of problems. In this chapter I will elaborate on the importance of understanding that subjects' performance in situations involving uncertainty is based on a theoretical framework that is different in important respects from any formal theory of probability. Additionally, I will argue that the outcome approach is reasonable given the nature of the decisions people face in a natural environment. To this end, I will review research which suggests some reasons why causal as opposed to statistical explanations of events are salient and functionally adaptive.

The primary goal of the outcome approach is to successfully predict outcomes in uncertain situations. When required, outcome-oriented individuals will provide values which gauge their confidence in the prediction. In this respect, the outcome approach is similar to a personalist theory in that probability is a measure of degree of belief. However, there are at least two major differences between the two theories. First, personalist theories were motivated by the desire to put subjective probabilities on a rational and scientific basis. Thus, among other requirements in these systems, subjective probabilities of repeated events should, over a long series of observed trials, closely approximate the actual frequencies of occurrence.

> If a person assesses the probability of a proposition being true as .7 and later finds that the proposition is false, that in itself does not invalidate the assessment. However, if a judge assigns .7 to 10,000 independent propositions, only 25 of which subsequently are found to be true, there is something wrong with the assessment.... The attribute that they lack is called calibration.... Formally, a judge is well calibrated if, over the long run, for all propositions assigned a given probability, the proportion of true equals the probability assigned. (Lichtenstein, Fischhoff, & Phillips, 1981, p. 2)

The outcome-oriented individual appears uninterested in calibration as defined above, but rather is interested in whether or not, on a particular occasion, a correct prediction can be made. If a prediction turns out to be in incorrect, the prediction was wrong and the confidence value, if assigned, was erroneous.

The other, and related, difference between the outcome and personalist approach is the method by which frequency data are translated into a confidence value. Since a goal in a personalist theory is to be well calibrated, frequency of past occurrence of some event is interpreted directly as a confidence in the future occurrence of the event, all other things being constant. (Applying Bayes' Theorem to subjective probabilities is one algorithm that guarantees good calibration.) In the outcome approach, frequency data are not directly used to formulate a probability. It is especially clear in the Cab, Painted-die, and Bone problems that frequency information is first translated into a more qualitative belief from which a numeric confidence is subsequently generated if necessary. A similar two-stage process of generating subjective probabilites has been suggested by Adams and Adams (1961) and more recently by Koriat, Lichtenstein and Fischhoff (1980).

> To assess one's confidence in the truth of a statement, one first arrives at a confidence judgment based on internal cues or "feelings of doubt".... The judgment is then transformed into a quantitative expression, such as a probability that the statement is correct. (Koriat et al., p. 108)

It should be added that the latter step of transforming internal cues into probability values is probably not an essential component of the outcome approach outside the laboratory. It seems to be done, and often begrudgingly, only if a request for a percentage or probability is made.

In the outcome approach, discriminating between small differences in the strength of these inner feelings is not what is most important. If the goal is to predict the most likely outcome on a particular occasion, one only need be aware of which outcome is associated with the strongest inner feeling. In this respect, it has been argued that people are well-prepared for the majority of decisions they face from day to day (Fhaner, 1977). When a decision must be made, seldom are frequency data available. Not only are they unavailable, but in many of the situations we face, frequency data would be of limited and questionable value. It is difficult to imagine, for example, what frequency data one might want in trying to make a decision about marriage, or of whether or not to start building a bomb shelter. For most decisions, however, frequency data could be of considerable help--Which car to buy, which college to attend, which lawyer to hire, which judge to hope for and which jury to select.

The benefits of simplifying such problems by employing judgment heuristics seem obvious, considering the physical or mental space which would be required so that relevant data could be held readily available. According to this "heuristic hypothesis," it is possible that at some more basic level, people reason in accord with formal probability theory, but to save time, heuristics are usually, and consciously, employed. In fact, people could be viewed as acting consistently with formal probability theory if they are believed to be using judgment heuristics because "they often work with a rate of success that is more than fortuitous" (Kahneman & Tversky, 1972, p. 452).

That people do not have conscious control over these "heuristic" judgments is suggested in the Bone and the Painted-die problems. When given data that required no summarizing, subjects still appeared to translate the data into more qualitative judgments from which probability estimates were derived. In the case of the Painted-die problem, this processing led to a prediction that was inconsistent with what one would expect on the basis of the representativeness heuristic.

In the outcome approach, the transformation of frequency data into beliefs is not accomplished via heuristics, but through causal analysis and theory construction. That is, the outcome-oriented individual contructs theories about the phenomena under consideration that are useful not only as a basis of prediction but, perhaps more importantly, as a basis of explanation.

The tendency to search for causal as opposed to statistical explanations was apparent in subjects' reponses to most of the problems used in this study. The importance of causality in making probabilistic judgments has been demonstrated in a variety of other contexts. Azjen (1977) and Tversky and Kahneman (1980) have suggested that base-rate information is more likely to be incorportated into probability estimates if presented in a way that strongly implies a causal link between features related to the data and the event of interest. Similar to performance on the Misfortune problem, subjects given biographies of deviants tend to reconstruct the information so that the plight of the "victim" can be viewed as an inevitable result of life-events (Rosenhan, 1973). Also, subjects given descriptions of accidents search for a pattern in the associated events that make the accident appear predictable and avoidable (Walster, 1962).

Several hypotheses can be advanced as to why people are prone to causal explanations even in the case of chance events (see Bulman, 1977). One claim, based on a Kantian epistomology, is that the way we construct reality presupposes causality and that the perception of causality, therefore, is adaptive in an evolutionary sense:

> ...attribution processes are to be understood, not only as a means of providing the individual with a veridical view of his world, but as a means of encouraging and maintaining his effective exercise of control in that world. The purpose of causal analysis--the function it serves for the species and the individual--is effective control. (Kelley, 1971, p. 22)

Another hypothesis is that in as much as a technological society is founded on the ability to exercise control over the environment, a causal view is

> ...connected with our manipulative techniques for producing results.... Thus a statement about the cause of something is very closely connected with a recipe for producing it or preventing it. (Gasking, 1955, p. 483)

A causal orientation is advantageous not only if it leads to a degree of control over one's environment, but also if it gives even the illusion of control, as suggested in the response of Subject 17 to the Misfortune problem:

> S17: I think that when people are able to predict--lend some consistency to events in the future--they feel more safe.... And I think that people could look into the future and have some sort of feelings about it, and it's a future that would even make them feel more at ease.

This subject acknowledged that his own chance explanation for the Misfortune problem would provide no such sense of security, but felt nevertheless resigned to his belief that co-occurrences of that type are more likely than not to be fortuitous.

Chance explanations are based on a model not only in which the notion of control is abandoned, but where predictability is often limited to aggregate as opposed to individual events.

> The separate throws of this series seem to occur in utter disorder; it is this disorder which causes our uncertainty about them. Sometimes head comes, sometimes tail comes, sometimes there is a repetition of the same face, sometimes not. So long as we confine our observations to a few throws at a time, the series seems to be simply chaotic. But when we consider the result of a long succession...a kind of order begins gradually to emerge.... (Venn, 1962, p. 5)

That subjects seem unable to understand, or at least to accept, non-causal models, can be viewed as a deep commitment on their part to the tenents espoused by the scientific enterprize of prediction and control. A similar commitment has been demonstrated by physicists in a refusal to accept the probabilistic foundation and implications of quantum mechanics (e.g., Bohm, 1957). Nor can the hypothesis that causal theories provide better predictability in most real-life situations be rejected. It is still an open question whether subjects' use of frequency data to formulate causal explanations typically results in better predictability than the direct assessment of readily available frequency data would provide. Whether or not people can predict more accurately using causal as opposed to statistical information, they appear to operate on the assumption that they can. This assumption is the heartbeat of the outcome approach.

References

Adams, J. K., & Adams, P. A. Realism of confidence judgments. Psychological Review, 1961, 68, 33-45.

Ajzen, I. Intuitive theories of events and the effects of base-rate information on prediction. Journal of Personality and Social Psychology, 1977, 35, 303-314.

Beach, L. R., & Scopp, T. S. Intuitive statistical inferences about variances. Seattle: L. R. Beach, 1967. (Mimeo)

Brown, D. L. Non-independence in subjectively random binary sequences (Research Bulletin No. 27) Princeton, N.J.: Educational Testing Service, 1964.

Bohm, D. Causality and chance in modern physics. New York: D. Van Nostrand, 1957.

Bulman, R. J. Causality: Psychological and philosophical perspectives. Unpublished doctoral dissertation, Northwestern University, 1977.

Carnap, R. Logical foundations of probability. Chicago: University of Chicago Press, 1950.

de Finetti, B. Foresight: Its logical laws, its subjective sources. In H. E. Kyburg Jr., & H. E. Smokler (Eds.), Studies in subjective probability. New York: Wiley, 1964.

Dempster, A. P. The theory of statistical inference: A critical analysis. Research report (S-2), Department of Statistics, Harvard University, 1968.

Edwards, W., & Kramer, E. F. Boodle from Bayes' at a penny a peek. Ann Arbor: W. Edwards, 1963. (Mimeo)

Edwards, W., Lindman, H., & Phillips, L. D. Emerging technologies for making decisions. In, New directions in psychology. II. New York: Holt, Rinehart & Winston, 1965.

Fhaner, S. Subjective probability and everyday life. Scandinavian Journal of Psychology, 1977, 18, 81-84.

Foucault, M. The order of things. New York: Vintage Books, 1973.

Gasking, D. Causation and recipes. Mind, 1955, 64, 479-487.

Hacking, I. Logic of statistical inference. London: Cambridge University Press, 1965.

Hacking, I. The emergence of probability. London: Cambridge University Press, 1975.

Jeffreys, H. Theory of Probability (2nd ed.). Oxford: Clarendon Press, 1948.

Jenkins, H. M., & Ward, W. C. The judgment of contingency between responses and outcomes. Psychological Monographs, 1965, 79, 1-594.

Kahneman, D., & Tversky, A. Subjective probability: A judgment of representativeness. Cognitive Psychology, 1972, 3, 430-454.

Kahneman, E., & Tversky, A. On the psychology of prediction. Psychological Review, 1973, 80, 237-251.

Kahneman, D., & Tversky, A. Variants of uncertainty. Cognition, 1982, 11, 143-157.

Kelley, H. H. Attribution in social interaction. Morristown, N.J.: General Learning Press, 1971.

Konold, C. E.,& Well, A. D. Analysis and reporting of interview data. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, 1981.

Koriat, A., Lichtenstein, S. & Fischhoff, B. Reasons for confidence. Journal of Experimental Psychology: Human Learning and Memory, 1980, 6, 107-118.

Kuhn, T.S. The structure of scientific revolutions. Chicago: University of Chicago Press, 1962.

Lichtenstein, S., Fishhoff, B., & Phillips, L. D. Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, and A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases. New York: Cambridge University Press, 1981.

Lucas, J. R. The concept of probability. Oxford: Clarendon Press, 1970.

Mackie, J. L. Truth, probability and paradox. London: Oxford University Press, 1973.

Magoon, A. J. Constructive approaches in educational research. Review of Educational Research, 1977, 47, 651-693.

Maistrov, L. E. Probability theory. New York: Academic Press, 1974.

Marton, F. Describing conceptions of the world around us. Reports from the Institute of Education, University of Goteborg, No. 66, 1978.

Mellor, D. H. The matter of chance. London: Cambridge University Press, 1971.

von Mises, R. Probability, statistics and truth. London: Allen and Unwin, 1957.

Nisbett, R., & Ross, L. Human inference: Strategies and shortcomings of social judgment. New Jersey: Prentice-Hall, 1980.

Peterson, C.R., & Beach, L. R. Probability learning? Ann Arbor: C. R. Peterson, 1967. (Mimeo)

Ramsey, F. P. Truth and probability. In R. B. Braithwaite (Ed.), The foundations of mathematics and other essays. London: Harcourt, Brace and Co., 1931.

Reichenbach, H. The theory of probability. Los Angeles: University of California Press, 1949.

Rosenhan, D. On being sane in insane places. Science, 1973, 79, 250-252.

Savage, L. J. The foundation of statistics. New York: Wiley, 1954.

Todhunter, I. A history of the mathematical theory of probability from the time of Pascal to that of Laplace. New York: Chelsea, 1949. (Originally published, 1865.)

Toulmin, S. E. The uses of argument. Cambridge: Cambridge University Press, 1958.

Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. Cognitive Psycholgoy, 1973, 5, 207-232.

Tversky, A., & Kahneman, D. Judgment under uncertainty. Heuristics and biases. Science, 1974, 185, 1124-1131.

Tversky, A., & Kahneman, D. Causal schemata in judgment under uncertainty. In M. Fishbein (Ed.), Progress in social psychology. Hillsdale, N.J.: Erlbaum, 1980.

Venn, J. The logic of chance (4th ed.). London: MacMillan, 1962. (Originally published, 1888.)

Walster, E. "Second guessing" important events. Human Relations, 1967, 20, 239-250.

Ward, W. C., & Jenkins, H. M. The display of information and the judgment of contingency. Canadian Journal of Psychology, 1965, 19, 231-241.

Well, A. D., Pollatsek, A., & Konold, C. E. Probability estimation and the use and neglect of base-rate information. Unpublished manuscript, University of Massachusetts, Amherst, 1981.

Wilson, S. The use of ethnographic techniques in educational research. Review of Educational Research, 1977, 47, 245-265.

APPENDIX A. TABLES

TABLE 1

SUBJECT BACKGROUND INFORMATION

Subject Number	Sex	Major	College Math & Statistics (Interview 1)	College Math & Statistics (Interview 2)
1	M	Psych	H.S. STAT	Intro STAT*
2	M	Psych	H.S. STAT	Intro STAT*
3	F	Psych	precalculus*	
4	M	Forestry	calculus Intro STAT	
5	M	Comm. Disorders		
6	F	Psych	precalculus*	Intro STAT*
7	M	Psych		precalculus*
8	F	Psych	precalculus Intro STAT	
9	F	Psych		
10	F	Psych	precalculus	
11	M	Psych	precalculus	
12	F	Elem. Ed.		
13	F	Pub. Relations		
14	F	Psych	Intro STAT*	
15	F	Psych	precalculus Business STAT	
16	F	Psych	Intro STAT	

*Currently enrolled

TABLE 3

EVALUATIVE RESPONSES AND YES/NO PREDICTIONS ON THE BONE AND WEATHER PROBLEMS

	Evaluative Response			Yes/No Predictions		
	Bone	Weather		Weather		
S#	I was right (wrong)	Forecaster was wrong	7/10 = pretty accurate	Compare value to 50/100%	Qualitative interpretation of 70%	50% = Anything can happen
1	x			x		
2						
3	x	x	x	x		
4			C*			
5	x		C	x	x	
6			x	x		
7						
8			C	x		x
9	x	x	x	x		x
10	x		x			
11	x	x	C			
12	x	x	x	x	x	
13	x	x	x	x	x	
14			x	x		x
15		x	x	x	x	
16	x		x	x		

*C indicates conflict between "pretty accurate" and "perfect."

TABLE 4

CAUSAL-FEATURE RESPONSES ON THE WEATHER, BONE, AND MISFORTUNE PROBLEMS

Subject No.	Weather: 70% = measure of causes	Weather: Error due to variability of weather	Bone: More trials of questionnable help	Bone: Ignore 1000 data in predicting 10 trials	Bone: Predictions based on physical features	Bone: The roll makes a difference	Misfortune: No mention of chance	Misfortune: Attribution of bad luck	Misfortune: External-controlling force	Misfortune: Internal-causal connection	Subject No.
1											1
2											2
3	x			x		x	x	x			3
4											4
5	x					x				x	5
6			x		x						6
7									x	x	7
8		x				x					8
9	x	x	x		x	x				x	9
10			x								10
11			x			x	x		x	x	11
12	x	x	x	x	x	x			x	x	12
13				x		x	x		x		13
14										x	14
15						x					15
16											16

TABLE 5

RESPONSES TO THE CAB PROBLEM

		Probability Estimate		"Deterministic" Responses (Outcome-oriented)		
Subject No.	Outcome score	Incongruent information	Congruent information	Asked if number wanted	Qualitative response	Estimate based loosely on data
1	3	.80	.80	x	x	
2	0	.80	.80			
3	12	.80-.90	.90-.95			x
5	8	.80	.80	x		
6	5	.60	.80-.85			
7	1	.80	.80			
8	5	.60	.80-.90			
11	9	.80	.85-.90	x		
12	15	.70	.70		x	x
13	12	.80	.90	x		x
15	7	.80	.85	x	x	
16	5	.80	.80			

TABLE 6

RESPONSES TO THE BONE-2 PROBLEM

		Desired Information						Basis of Prediction		Statistician's Estimates	
Subject No.	Outcome score	Surface area	100 trials	1000 trials	Center of gravity	Bone to look at	Ss' last 10 trials	Frequencies	Physical Properties	Frequencies	Physical Properties
1	3		2*	1				x		x	
2	0		2	1				x		x	
3	12					2	1	x	x		x
5	8			1		2		x	x	x	x
6	5		2	1				x		x	
7	1			1		2		x		x	
8	5		2	1				x		x	
11	9		2	1				x		x	x
12	15	2				1			x	x	x
13	12		2	1				x		x	x
15	7			1		2		x		x	x
16	5		2	1				x		x	

*1 indicates 1st choice; 2, 2nd choice.

TABLE 7

RESPONSES TO THE PAINTED-DIE PROBLEM

Subject No.	Outcome score	Expected No. of White: In six trials	Expected No. of White: In sixty trials	Expected No. of White: Average trials to get white	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6
1	3	1	10		5/6B (W)	· (B)*	· (B)	· (W)	· (B)	· (B)
2	0	1	10		5/6B (B)	· (B)	· (B)	· (B)	· (B)	· (W)
3	12	0		6-7	.90-.95B (B)	· (B)	.85-.90B (B)	.80B (W)	.90B (W)	· (B)
5	8	1	10		5/6B (W)	(B)	· (W)	· (B)	· (B)	· (B)
6	5	1	10-12		1/6B (B)	· (B)	1/6W (B)	1/6B(B)	1/6W(B)	1/6(B)
7	1	0	<10-10		5/6B (B)	· (B)	· (B)	· (B)	· (B)	· (B)
8	5	1	10		.80B (B)	· (B)	· (B)	· (B)	· (B)	· (W)
11	9	0	10		5/6B (B)	· (B)	· (B)	· (B)	· (B)	· (B)
12	15	0	4		.90B (B)	· (B)	· (W)	· (B)	· (B)	· (W)
13	12	1*		10	.90B (B)	· (B)	· (B)	.10W(B)	.90B(W)	· (B)
15	7	0	5		.95B (B)	· (B)	· (B)	· (B)	· (B)	· (B)
16	5	0		12	5/6B (B)	· (B)	· (B)	· (B)	· (W)	· (B)

*1 for stones; 0 for die.

*dot indicates same prediction as previous trial.

TABLE 8

RESPONSES TO THE MODELLING PROBLEM

Subject No.	Outcome score	Die: Stones not model of die	Bone: Model not spontaneously developed	Bone: Rejected urn-model based on statisticians' estimates	Bone: No model possible
1	3				
2	0				
3	12	x	x	x	x
5	8				
6	5		x	x	x
7	1		x	x	
8	5				
11	9				
12	15		x	x	
13	12	x	x	x	x
15	7		x	x	
16	5				

APPENDIX B. INTERVIEW 1 TRANSCRIPT

Subject: S11
Interviewer: C. Konold
Date: Dec. 1981

Weather problem

001 I: First, what does it mean when a weather forecaster says "tomorrow there's a 70% chance of rain"?

002 S: It might rain.

003 I: Uh-huh.

004 S: Other than that--It's weird because I don't really pay attention to weather forecasts, so

005 I: Uh-huh.

006 S: It's a different situation, I'll tell ya. Given how I pay attention to weather forecasts, to me it means pretty much nothing.

007 I: Uh-huh.

008 S: You know, I mean, I guess to some people you know it's important to them. I mean er, it's weird 'cause--yesterday--yesterday someone was talking about a er, they were upset--or not yesterday the day before--they were upset by the--about the snowstorm that we had because the weatherman hadn't predicted it.

009 I: Uh-huh.

010 S: And so I guess some people guide you know, or guide, to to some extent, their behavior on it but, I don't know. It really doesn't mean too much to me, I don't know.

011 I: What does the number--in this case, the 70%--tell you?

012 S: Oh, well, it's a high probability or likelihood that it will rain.

013 I: Uh-huh. If you'd--Go ahead.

014 S: I'm not sure--like in what context do you mean--you know, "What does it mean"?

015 I: I'm just wondering in this particular context, if you heard someone say there's 70% chance of rain tomorrow, what that number 70% would tell you. For example, suppose you heard instead, it was 50% chance of rain tomorrow.

016 S: MmMm. Ok, I see what you're saying. So, well I guess the 70% would mean that it's, you know, it's higher likelihood that it would rain, as opposed to the 50. I don't know. I just uh, with the weather forecasting it's not--your know, It's not--It's like it can change.

017 I: Uh-huh.

018 S: so quickly that the number itself doesn't seem as though it's that much of a clue to whether or not it's going to rain.

019 I: Any idea of where they come up with that number, you know, the 70%?

020 S: Where the weather people would do it?

021 I: MmMm.

022 S: Um, well let's see. I have a friend who is a meteorologist and so she should probably, you know, in my talking with her--just from conditions, I guess, you know, from atmosphere and from var--things--variables, I guess. Yeah. Uh, the--I guess the weather is just like a lot of problems where you're looking at a whole bunch of variables and as they, as they come to gether in one situation, you know, how they're--and today, you know, if you look out and you see clouds here and they're so high up, and there's moisture here or whatever. Then it kind of adds up to a likelihood that it will do something and maybe the 70% is--maybe when they say that it has to do with how sure they are that it's going to happen, I don't know.

023 I: Suppose a forecaster says that there is a 70% chance of rain tommorrow and you went out the next day and in fact it didn't rain. What would you conclude about that statement that had been made, that there was a 70% chance of rain?

024 S: Well, this--probably what would be most important uh, if it didn't rain, umm, would be in how I would assess, you know, the next statement. You know, if they said that the next day it was going to rain.

025 I: Uh.huh.

026 S: And even if it was 90% I would be carrying with me, you know, the idea that he or she or whatever--it was predicted that it was going to rain the next day with a 70% chance and he was wrong or they were wrong. And now, you know, I'll take into consideration on the next occasion

027 I: Uh-huh.

028 S: uh, that they were wrong on one occasion. And that wasn't very well said. But you kind of get the idea; yeah. It would just kind of help me, you know, decide on whether or not I think it's going to rain the next time they say it's going to rain.

029 I: Suppose you wanted to determine how accurate a particular forecaster was

030 S: MmMm.

031 I: in making certain predictions. And on 10 days, you kept track on 10 days when they predicted 70%

032 S: right.

033 I: Um, chance that there would be rain. And on those 10 days that they predicted that there was 70% chance of rain, um, on 3 of them, in fact, it didn't rain. What would you conclude about the accuracy of this particular forecaster?

034 S: Oh, it seems as though they were pretty much right on. Because, you know, in the whole of the 10 days they were right, uh, on three of the occasions, which would be 30%, they--it didn't rain and that's what they kind of predicted, almost, maybe.

035 I: Uh-huh. What are you thinking?

036 S: Well, cause it's like--they didn't predict a 10 day span: They predicted each individual day.

037 I: Uh-huh.

038 S: And so it seems as though when each day comes up it's a whole in itself, it's not necessarily put together in a unit.

039 I: Uh-huh.

040 S: So that--I guess it's almost like when you roll--it's like, I guess, when you roll the dice. You could say that uh, or when you flip a coin or whatever, each time you do it, you have the same chances of it happening.

041 I: Uh-huh.

042 S: And so it would, let's see. Yeah. So each time you do it it's the same chance that it's going to happen, that it's either going to rain or not rain 70%. But then in the whole, as you add all the different times that you do it together, it almost seems that it's not the same.

043 I: It's not the same?

044 S: Yeah, it's not the same. It's almost like you can't add them together. But maybe you can, I don't know. This is hard, I don't really understand.

045 I: Maybe you can just explore why that seems inappropriate to add them together.

046 S: Mmm.

047 I: You mean--what you're saying somewhat is to look at their accuracy over 10 days is not quite appropriate? I mean, because they're predicting for each individual day?

048 S: Right. And yeah, so it kind of gets back to the idea that they were pretty much right getting 7 out of 10 right.

049 I: Uh-huh.

050 S: Umm, maybe it's not fair to, to judge it that way. Maybe you should just judge it on each one. But I guess--well maybe you can add them together beacuse they're all the same. They're all like 70% and um, no, it really can't rain 70% of the--well, can it rain 70% of the day? Is that what they mean? Or is--I don't know. Er, it would depend on whether or not it rained. It would first depend, I guess, how you would quantify if it rains.

051 I: Uh-huh.

052 S: And then if it rains on the day, or if you just have to have a few drops as opposed to it raining for a period of time--maybe that would affect it.

053 I: Uh-huh.

054 S: This is a logic swirl.

055 I: So you're a little bit uncertain about what the accuracy is or how to assess the accuracy.

056 S: Right. Yeah. Starting with the definition of the 70% accuracy or 70% of the raining for the day.

057 I: So I guess my question--how accurate was this forecaster? He missed on 3 of the 7 days in predicting rain--

058 S: MmMm.

059 I: What's your er, you're a little confused about that. What's your sense about how accurate this forecaster is?

060 S: Well, I probably would say that he was pretty close because

061 I: MmMm.

062 S: he was just--because he thought, or she thought um, it was going to rain on--it had a 70% possibility for each of the days and he was right 7 out of 10 times, or she was right, and that's, I don't know, it just--

063 I: Would they have been more accurate had it rained on all 10 of the days? Would that be more impressive to you?

064 S: Well that's weird 'cause--'cause it almost seems that it's--it almost seems, I don't know.

065 I: What are you thinking?

066 S: Well, I'm thinking that if they go over, it's almost like they're going over, and they're wrong the other way.

067 I: Uh-huh.

068 S: It's raining more than they really predicted.

069 I: Uh-huh.

070 S: But it's not like they predicted that it would rain 70% on 70% of the days of these next 10 days. It's like they predicted rain for each day.

071 I: Uh-huh.

072 S: And if--and they were 70% sure, then they--then if they had a--And if they were 70% sure that it was going to rain on each day, and then as the higher--yeah. If they had more--if it rained on more of those days it would seem as though they, um, that they were more accurate.

073 I: MmMm.

074 S: But then if it's only supposed to rain on 70% of those days then as you increase the number above 70-7, then they would be less accurate.

075 I: So you're sort of in a dilemma there?

076 S: Yeah, I don't know, it depends.

077 I: Umm, numbers are often assigned to various events like elections, sporting events, causes of death. For example you might hear something like er, someone will make a statement like "the equal rights ammendment has a 25% chance of being ratified by the June '82 deadline."

078 S: MmMm.

079 I: Um, if you heard a statement like that, what would that 25% in that particular context mean to you?

080 S: It doesn't sound good.

081 I: Uh-huh.

082 S: You know. I guess it kind of clues you in on whether or not it's going to pass or if it is going to be ratified. Um, and, but shoot. I don't know what I should do with it past that.

083 I: Any idea where that number would've come from? When you hear that 25%, what does--25% of what?

084 S: Yeah. I don't know. It would be really, I think that would be--If it was like in a context of every day, I would probably just, like, I would think of it as meaningless.

085 I: MmMm.

086 S: Umm, yeah, I don't know, because, uh, probably one of the reasons is, is 'cause I had a um, an intro methods course

for sociology. And some of the statements that we were for--not forced to make but that we had to make, were just very, er, you know--using numbers and trying to be persuasive with numbers. They weren't--it wasn't useful. You know.

087 I: Can you give an example?

088 S: Umm, it's, well, I'm thinking of elections because I took it during the election, while the 1980 election was going on--the presidential election--and we did something with the--like, I don't know, Carter had such and such a percentage of getting a certain number of votes.

089 I: MmMm.

090 S: And, Uhh, I don't know how we got it. I can't remember how we got the percentage. But it was kind--I don't know, it wasn't--it didn't really have any meaning to me because--Some people will look at that, like they'll--I guess one of the complaints about having the live coverage of the big campaigns is that some people will be persuaded at the last moment--moment, given who's running and how much percentage they're getting of the uh, of the incoming ballots, and then it will influence their voting behavior. And so I guess what I kind of meant, uh, that it wasn't useful is that, uh, it didn't clue you in on anything about the situation. It didn't give you any information about whether or not, for example, the E.R.A.--it doesn't tell you whether or not it should be ratified. It's just, you know, it's just kind of like a number that's thrown at you that it might be.

091 I: Uh-huh. There's another, er, you hear a statement too, like a person's chance of dying before the age of 40 are 25%. What does that 25% in that case tell you?

092 S: Mmm, mortality rate; how many prople are dying,

093 I: MmMm.

094 S: maybe, before that.

095 I: Is that easier for you to understand?

096 S: Yeah. Kind of. Yeah, it seems like it's easier to pigeon hole it--It's easier to, er, give--to figure out--or to look back on what kinds of things that they would be looking at, you know.

Bone problem

001 I: Um, I have a bone here that has 6 surfaces and I've written the letters A through F, um, one on each surface. If you were to roll that, which side do you think would land upright? You can sort of inspect it but don't roll it, yet.

002 S: Don't roll it? OK. Should I try to figure out--verbally?

003 I: Sure: Yeah, I'd like that.

004 S: Ac--what I'm trying to do is, um, is just, yeah; figure out which side would land and stay on the bottom.

005 I: MmMm.

006 S: And so--it would almost seem as though the flatter end would try to stay on the bottom, I would think, you know, if you were to roll it.

007 I: MmMm.

008 S: I'm almost trying to picture it rolling. Um, well, and I guess I'm assuming that--that it will, er, that there is, uh, there is a preferred orientation like if you do roll it ,

009 I: MmMm.

010 S: you know, rather than with a dice that's fixed. Um, which is, I guess, symmetrical, and if you roll it it'll--can pop up anywhere. I guess I think I would go with the flatter end or side would be on the bottom, and, yeah.

011 I: MmMm. So which side do you think would land upright?

012 S: Yeah, so I think probably C would land up. Right.

013 I: MmMm. OK. And how likely is it, do you think that C would land upright?

014 S: Well, trying to assess that, I would probably go the next step and figure out, try to figure out which would be the next, if any--which I would--which side I would think would come up next and then kind of--

015 I: MmMm.

016 S: And if I wanted to try and be accurate, I'd try to figure out the 3rd one,

017 I: MmMm.

018 S: and then try to draw some type of a, um, some type of a relationship between the 1st, 2nd, and 3rd, and if I really wanted to be deep, I'd probably look at every single side.

019 I: MmMm. Well, maybe without being deep you could just give an estimate--a real quick estimate.

020 S: Well, figuring that this is pretty big--it's got this point here, that would point--well stop it.

021 I: MmMm.

022 S: Umm, Oh, 1, 2, 3, 4, 5--6 sides.

023 I: There are 6 only.

024 S: 6 sides. So it's got 1 in 6 chance of coming up.

025 I: What, er, what did you say? One in--

026 S: 1 in 6 chance of coming up, if it was just,

027 I: MmMm.

028 S: you know, if it would come up. If each side has an equal chance of coming up, then it--1 in 6.

029 I: MmMm.

030 S: But, uh, since I think it's a preferred side, make it 1 in 3, umm, and, of course this side--1 in 3. So that would be like 33 or 30%, or something like that.

031 I: Why don't you go ahead and roll it and see what happens.

032 S: [Rolls C] Oh, wow!

033 I: Wow what? Hold it before you roll it anymore.

034 S: How come that came up just like that?

035 I: What do you mean?--That's--

036 S: I can't believe it.

037 I: You're surprised that it came up the way you predicted?

038 S: Yeah. That's weird.

039 I: How come?

040 S: Well, because, uh, I only predicted 1 in 3 and

041 I: Uh-huh.

042 S: baby came up first time. It's not supposed to do that. It's supposed to be like, um, you roll the first two and it's gotta, you know, make it on the last one.

043 I: Uh-huh.

044 S: And then it makes it--It's a more fun game that way.

045 I: So that was weird?

046 S: Yeah, it was. Yeah, that's pretty strange.

047 I: What do you conclude about your prediction?

048 S: I guess it was right.

049 I: Uh-huh.

050 S: So I guess I looked at the right things or whatever.

051 I: So what do you conclude now, having rolled it once?

052 S: That, uh, maybe it will come up more than, you know--so maybe it will come up more than 1 in 3.

053 I: Would rolling it more help you conclude which side--

054 S: It would seem as though it would. Yeah. It seems a little higher, yeah, given the higher n, you know.

055 I: Why don't you roll it as many times as you like and sort of let me know what you're thinking or concluding as you go along.

056 S: [Rolls A] Allright. That's a whole opposite side.

057 I: That's A.

058 S: Well that's good. That's better. [Rolls B] Alright, so maybe it isn't.

059 I: Why are you saying 'alright'?

060 S: Well, 'cause it seems as though if there are--if a bunch of them are coming up--so maybe it's a real--you could just go out and find these bones, if you don't want to buy die.

061 I: Uh-huh.

062 S: You can have a fun game of playing bones instead of die.

063 I: So you're thinking it's--

064 S: Maybe it's, uh, maybe, um, it'll come up any way.

065 I: MmMm.

067 S: You know, so maybe it doesn't have a preferred orientation. That's what I guess I was thinking. I'll try it again. [Rolls B].

068 I: B.

069 S: Yep. So what I would probably do as I was continuing to roll it, you know, if I wanted to figure out if it was, er, if it did have like a preferred orientation

070 I: MmMm.

071 S: is I'd probably just add them up, you know, in 6 columns

072 I: MmMm.

073 S: and see which ones came up the most. And I guess E and F don't seem to be coming up. Maybe because it's too heavy. That would be pretty--Hmm.

074 I: So how many times would you, er, if you kept track, how many times would you want to roll it before you concluded which side?

075 S: Before it was a pretty certain

076 I: MmMm.

077 S: estimate? Umm, I don't know. Maybe, give it a--you have to give it enough of a chance. Uh, probably like a couple each. Well, so that would be 2, 4--something like 12 or 15 or--it wouldn't, er, I guess you probably could do it in a short, short enough time--I mean number of rolls.

078 I: MmMm. You want to roll it a few more times and see what you think?

079 S: [Rolls D] So D comes up. [Rolls C] C comes up. [Rolls B] So that--B has three. [Rolls B] Four. [Inspects bone] So maybe it's that bump that throws it off. [Rolls C, D] A doesn't seem to be coming up. It seems to be B, C, and D. [Rolls C, C] Hmm, I don't know. It's--It's like, well it seems that what I'm getting is that it does prefer, um, this--you know, these, uh, and C er, the letters in this orientation.

080 I: MmMm.

081 S: But it doesn't--it seems to prefer uh, B, C and D.

082 I: MmMm.

083 S: [Rolls D] But it hasn't landed that much on A.

084 I: Do you have any idea of which of those 3, which is most likely?

085 S: Well, I'll go for that one, [as rolling--C]. Oh, C. So C is going to be it. But I think it's B.

086 I: You think it's B?

087 S: Yeah, I think it's B. Yeah, but it's real close. So that if I was to give a, you know, like a--to give a percentage, it would probably be like uh, let's see, 40, 30.

088 I: 40 for B?

089 S: Yeah, 30 and 30.

090 I: 30 for C and 30 for D?

091 S: Yeah. Something like that. But except that then well, you'd have to take into account that these would come up once in a while, E and F and A would come up once in a while. So I guess it would be like--

092 I: Have you got any A's, do you remember?

093 S: I don't--no, I don't remember but it doesn't--maybe we got 1 or 2 but it doesn't seem--well, probably not 2. Then we'd probably would've remembered it.

094 I: MmMm.

095 S: It seems as though we got mostly B's and then C and D were about the same. Do you know what kind of bone this is?

096 I: Uh, it's from a deer; it's an ankle bone from a deer. Um, how many times would you want to roll before you were absolutely sure of what numbers to give to the sides?

097 S: Umm, absolutely sure? Uh, that's a weird one, 'cause, you know, it's like you, how do you, you know , where do you start making--what do you use to figure it out?

098 I: MmMm. Well just an estimate of how many times you'd roll and keep track of before you felt pretty confident about it.

099 S: Uh-huh. Probably, I don't know, something like--it would be boring, but something like 100.

100 I: MmMm.

101 S: You know, it would take a pretty high, you know--because it would be the first time that I've ever seen it, so, I mean--

102 I: I got real bored one day and rolled it a thousand times.

103 S: Wow!

104 I: And this is what I recorded.

105 S: So D is the preferred, huh? Wow! E came up? That's pretty weird [inspects bone].

106 I: So what--what are you thinking?

107 S: Oh, umm, well I would have thought that A would've come up more than E.

108 I: Uh-huh.

109 S: But it is kind of like--2 is pretty insignificant given a thousand.

110 I: Uh-huh.

111 S: But--And these two are pretty close.

112 I: B and C?

113 S: Yeah.

114 I: Would you be willing to conclude that B is more likely than C?

115 S: Yeah, but n-not it's--because the numbers are so close it's not--it's not that big a difference.

116 I: Would you be willing to conclude that D is more likely than B?

117 S: Yeah, that would--Yeah. That would be pretty--there's enough of a difference that it's definitely--sure.

118 I: If I ask you what's the chance of rolling a D, what would you say?

119 S: Oh, I'd almost say 1 in 4 because it's pretty close to 400.

120 I: MmMm.

121 S: And, so. Well, would it be 1 in 4? No, it would be 4 out of 10, whatever that would be. So a 40% chance.

122 I: OK. I'm going to ask you to roll it 10 times--the bone 10 more times, but before you do it, to predict how many of each letter you will get. Umm, and maybe you could write it up there and we could keep track.

123 S: OK. I think--it's alright if I just use this information here?

124 I: You can use whatever information you want to.

125 S: OK. This seems like it's more valid because it's got such a high n.

126 I: MmMm.

127 S: So what should I do, write on here?

128 I: Just write the letter and beside it, how many you predict you'll get out of 10 rolls.

129 S: OK. Um, I think I'm probably just going to go with B, C and D. Is that too big or--[Referring to the size of his writing]

130 I: No, that's fine.

131 S: OK. So it'd be out of 10, huh? That's going to take a little--let's see.

132 I: What are you trying to do?

133 S: Uh, well just to round up and round out--You know, round off these percentages--what would be percentages if you divided. Umm, well, let's see 3, 4, 5, 6, 7. I think--cause I have--at first I'm starting out with the uh, the 100's unit.

134 I: Uh-huh.

135 S: And I'm just gonna assign 3, 2 and 2, and then divide the last two or last three. And this would have to get 1.

136 I: The D would have to get 1 more?

137 S: One more. So it'd have to have at least 4.

138 I: MmMm.

139 S: And I would probably assign, I guess--let's see, maybe a second one and then one to here, yeah. So that would be a 5, 2, and 3.

140 I: MmMm.

141 S: Yeah. I guess that's good.

142 I: And why have you given none to either A or E?

143 S: Uh, 'cause they just--they have--if they--E and F didn't come up when I rolled it.

144 I: Uh-huh.

145 S: And it's really small--and A is smaller than E. And I don't remember if A came up but--

146 I: So it's somewhat based on what you remember rolling too?

147 S: Yeah, somewhat. Right.

148 I: OK. Well why don't you go ahead and roll that and keep track of what we get.

149 S: [Rolls D] So we have a D. Should I write it here?

150 I: Yeah. If you could just keep track or I'll forget.

151 S: OK. So we have one. [Rolls D] Oh, wow. Another D. [Rolls B] OK. We got a B. That's pretty good.

152 I: Why'd you say that's pretty good?

153 S: It's kind of staying in line, almost. [Rolls E] Oh, no. How did we get it? Oh, no! That totally throws it off. So we got to put it on this side.

154 I: Just write E and keep track there.

155 S: E--he got one. [Rolls B] Oh boy. Come on D.

156 I: Come on D, Huh?

157 S: Yeah, well, we want a-- [Rolls D] oh go--yeah. How come the C is not coming up though? So we got 1, 2, 3, 4, 5--

158 I: 6 so far.

159 S: [Rolls D, D].

160 I: That's 8 rolls. What do you predict you'll get on the next 2, here?

161 S: Predict? Probably D's. D's and a B, probably, well, Yeah. Eight--'cause, well--See, I have like a--you can look at it as 10 rolls or you can look at it as 1 roll and then 1 more roll. And--

163 I: What do you mean you can look at it as 10 or 1 roll and 1 more roll?

164 S: It's, well, I shouldn't really let this influence how it's going to roll again.

165 I: Uh-huh

166 S: 'Cause it's probably going to come up more D's--It has a higher chance of coming up D.

167 I: Uh-huh.

168 S: than anything else. But, I don't know, it's weird. It seems as though over the long run it should come up a certain--it should fall in line with that.

169 I: Uh-huh.

170 S: And maybe if you do have a high enough number of rolls. But it--each time you roll it, D is more likely to come up than anything else.

171 I: So is that why you're predicting a D and a B in the next two rolls?

172 S: Yeah. A D or B. But then, there's another part of me that wants to say that C will come up because it hasn't come up,

173 I: MmMm.

174 S: you know, which maybe is kind of false.

175 I: MmMm.

176 S: But--'cause there's no--it's not like--it's not like--this information isnt, kind of, willing over this problem [pointing to results of eight rolls].

177 I: MmMm.

178 S: It's gonna roll--that bone is going to do whatever it will do each time.

179 I: Mmm.

180 S: And so past isn't going to influence how it's going to roll in this one--I think that makes sense.

181 I: OK. Why don't we roll it and see what happens, here.

182 S: [Rolls C] What is it with this? An intelligent being here? That's what it is. I don't believe it.

183 I: So what do you think now?

184 S: I think it's a prop--(laughing) I'm just kidding. Well I still think that--what I said a minute ago, was true.

185 I: MmMm.

186 S: And C just came up because of the way it was rolled. [Rolls D] And that would kind of like further--that would back what I--it would seem to.

187 I: So how do you feel about your prediction?

188 S: Pretty awful. Well, actually, it wasn't too bad because uh, the E is only off by 1--er, the E--the D is off by 1. This [points to C] is only off by 1. Wait a minute. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. Go man, yeah. These [points to B and C] are only off by 1. This is off by 1, and this [points to E].

189 I: MmMm

190 S: That's pretty good. Sure.

191 I: If you were going to roll it 10 more times, and you were going to predict what you'd get, would you change what you--

192 S: I'd probably still go with this.

193 I: Uh-huh.

194 S: This is a freak. This is just--That will never happen again.

195 I: You won't ever get an E again?

196 S: Maybe, I don't know.

197 I: OK.

Misfortune problem

001 I: I know of a person to whom all of the following things happened on the same day: First, his son totalled the family car and was seriously injured; next he was late to work and nearly got fired; then he got food poisoning at a restaurant in the afternoon; and in the evening he found out that his father had died. How do you account for all those things happening on the same day?

002 S: Bad Karma.

003 I: Bad what?

004 S: Karma. You know, he just like--it wasn't--

005 I: Oh, bad Karma.

006 S: You know, it just wasn't happening for him that day.

007 I: Uh-huh.

008 S: But well, each one of those is a separate event and why they would happen on the same day, I don't know. There's no--I'm trying to look for clues, you know, trying to figure out why.

009 I: Is that a question that you'd ask yourself if you just heard that or that happened to a friend of yours?

010 S: Well, how I would react is just uh, if it was a friend of mine you know, probably the first--I wouldn't--I'd try to be--I'd try to just help him--

011 I: Uh-huh.

012 S: deal with it. Uh, but, and I don't know if I would--It would be kind of unusual. And so you would probably wonder "why me," you know, or "why him"--

013 I: MmMm.

014 S: or "why on the same day" uh, and maybe what I would do is take each one and--like for example with the--I guess the father died or something?

015 I: Right. First his son got in an accident and totalled the car. And next he was late to work and nearly got fired. And then he got food poisoning. And then he heard that his father died.

016 S: Yeah. I would probably just look at each one individually. And so if the son was like--I don't know--try to make each one fit. If his son was just a bad driver or whatever. Or if he was a good and just--it was, you know--or he'd never gotten into an accident or something. Try to figure out something about--maybe, maybe just he'd had it good for too long and there was just like uh, all these bad things piling up, just waiting to come at him, or something.

017 I: MmMm.

018 S: I don't--I just try to look the--yeah. I guess I would be looking at the past to figure out--to give me, you know, an idea about what's, you know, why they kind of happened all at the same time.

019 I: Uh-huh.

020 S: But other than that, I don't know.

021 I: Have you ever had anything very unlikely happen to you?

022 S: Very unlikely?

023 I: Yeah.

024 S: Umm.

025 I: Other than rolling that C a minute ago.

026 S: Or the C the first time. I don't know. Let me see. Oh, sure. Oh, gee, yeah.

027 I: How did you make sense of that?

028 S: Um, well, it was a bad car.

029 I: Uh-huh.

030 S: It's just a lousy car. And uh, it was something like we were uh, I'm trying to remember the event. Oh yeah. OK. My car, it was just like back 5 or 6 years ago, my car had a problem with it's starter or something. And it wouldn't--sometimes you'd turn the key and nothing would happen. And you'd figured the battery was dead. And it wasn't the battery, and you'd figure something else; and it wasn't that. And it happened once or twice and then uh, and I'd curse at it. And a couple of minutes later it would start up. And I couldn't understand it. And then one night, we, er, it was during the summer and we had gone out to uh--a bunch of my friends and I, and we'd gone out to some amusement park.

031 I: Mmm.

032 S: And we were coming back and it was like 1 or 2 in the morning. And, you know we were in no condition to deal with a car that was, er, you know, that would die on us. And we went to a store and the car died on us. It was a really bad thing to have happen to us. We were like--we were up

in Boston and I lived, you know, on the Cape and it's a long walk home and--But, I think--and something else happened that night too, that was weird. 'Cause at the amusement park something happened: Oh, yeah. At the amusement park I hit someone. And uh, I'd never hit anybody, you know, in my life, you know.

033 I: Uh-huh.

034 S: And I wasted my front light on one side.

035 I: Uh-huh.

036 S: And let's see. Did anything else happen that day? Oh, that's right. 'Cause I was working at the time. It's weird because I was working at a motel with my sister--or either I was working there or I'd just started working there or I hadn't worked there or something. But, it was a job that my sister had gotten for me. And I was--I think I'd just started working there 'cause I didn't know the boss that well and I had to miss work that day--the next day--because I had to spend time at a friend's house because I didn't want to walk all the way to the Cape. And uh, I don't know. I guess er, so something kind of similar to those--

037 I: Uh-huh.

038 S: Those 3 events did happen.

039 I: Can you remember how, er, how you made sense of that?

040 S: Well, umm, I guess it's kind of--I guess a lot of them were my--I kind of put the blame on me in that uh, I never bothered to deal with the car. You know. I knew that it would weird out and I never dealt with it. And then when it just totally broke down, so that--I put that into--I tried to make a scheme for why each individual thing happened and then I hit the car because I wasn't looking or something. I tried to make it all fit is what I--But, you know, almost losing my job was, er, that was the shakiest part, because the job was new and it was a chance to, you know--

041 I: That almost sounds like this one.

042 S: Yeah.

043 I: Yeah, it was pretty bad. But I'd forgotten about that one.

It's a good thing too.

Stone problem

001 I: Let me move on here. Uh, in one of my pockets I have 2 stones and 1 is black and 1 is white. And in the other pocket I have 4 stones, 3 are black and 1 is white. I'm going to reach into

002 S: MmMm.

003 I: one of my pockets, which is always tough to do sitting down, and pull out a stone. What color do you think it is?

004 S: Well, is it pocket A or B?

005 I: That would make it easier but I'm not going to--

006 S: You're not going to say. Um, I'd probably say black.

007 I: MmM. OK. And why would you say black?

008 S: Because there's 5 er, 4 of them and there's only two white.

009 I: MmMm.

010 S: But if it was pocket A, it would be a tough decision. But I'll go with the easy decision.

011 I: The easy decision. And how likely is it that, in fact, it's white?

012 S: Oh, it's not--somehow, uh, let's see, 1, 2, 3, 4, 5, 6. Well, 1 in 3, maybe.

013 I: 1 in 3. How did you get 1 in 3?

014 S: I just added up the total number of uh, stones.

015 I: When you say "maybe," are you not quite sure if that's accurate.

016 S: Ya. It's not accurate.

017 I: How do you know it's not accurate?

018 S: That's one pocket. And that's another pocket.

019 I: Uh-huh.

020 S: And so given whatever pocket you--of course, um, yeah. Because you're selecting from 2 different uh, 2 different samples.

021 I: Uh-huh.

022 S: And--or samples or whatever, uh, and the probability is not equally likely that the white will come up in this situation as that.

023 I: MmMm.

024 S: Uh, and then--and also by reaching into just one of the pockets, you're not really reaching, or you're not considering all of these

025 I: Mmm.

026 S: together. You're only considering either this situation or this situation.

027 I: So are you saying the 1 in 3--

028 S: Is just er, just kind of trying to give an answer as opposed to really trying to give a very accurate answer.

029 I: Do you have any idea of how you would give an accurate answer? I'm not asking you to. But would there be a way to do it?

030 S: Umm, dealing with these things right there? No, probably something, uh about, wh--I'd just try and figure out which pocket you were going into, you know, reaching into, but I can't really--

031 I: Well, it, in fact, is black.

032 S: Hmm. Well, it was right.

033 I: Umm, which pocket do you think I drew that out of?

034 S: Probably B. And I would use it because there's 3 of them in that pocket and--

035 I: How likely is it that, in fact, I pulled it out of pocket A?

036 S: How likely? That's a tough one. I'm not sure what to consider in that situation. Umm, how likely--I don't know what to --

037 I: Maybe just give an estimate.

038 S: Well, it's weird because I'm not sure even what to consider.

039 I: Mmm.

040 S: It's like a--I guess you could just, er, like omit the whites and then consider just these 4. And since--well no. I can't do that. Uhh, 'cause if I don't consider these whites then it doesn't matter. They're all black in each case. So no matter what pocket you pull out you're going to get a black, but--um, so that, oh, I guess what I would do, maybe just--is try to figure out which pocket has--the black has the higher probability of coming out of.

041 I: MmMm.

042 S: And so the--is it a, er, the question is which pocket you pulled it out of?

043 I: Well, you said probably pocket B and I was asking you how likely you thought it was, that it, in fact, came out of pocket A.

044 S: Oh, how like, er--

045 I: Or you can answer the question how likely is it that it came out of pocket B--either question.

046 S: B? Oh, OK. So give a number to it. Um, well it would seem as though it would be 3 times as great, 'cause, yeah. Because we have 3 in this pocket so yeah. The likelihood would be--no. I don't know how to translate it into percentages, but it's 3 times.

047 I: 3 times as great?

048 S: Yeah.

049 I: I'm going to reach into the same pocket now and pull out another stone. What color do you think it will be?

050 S: Wow. Well, it's--well in the 2 situations--I would cover up both blacks,

051 I: Uh-huh

052 S: and then if it was this one, it'd be white. Sure. And if it was this one, it'd be almost close to getting white. So--and since I don't know which pocket it is, I would probably say wh--I would say white.

053 I: It'd be more likely to be white?

054 S: Yeah. Um, and I think--yeah. 'Cause I'd yeah. I would try to consider each one.

055 I: Uh-huh.

056 S: This is--that one--If that black is gone, it's gonna be white: If this black is gone, you know, you're increasing the chance that white's gonna come up. So--

057 I: So any--how likely do you think it is that it will be white?

058 S: Umm, well if it's the same pocket--So it's really likely here. It's like, it's, uh--And here it's pretty likely. And maybe it's not fair to draw--to add them together. Uh, so 1 in--well, no. A 50/50 chance, I'd guess.

059 I: 50/50. How did you get 50/50?

060 S: I added together.

061 I: Just the two whites?

062 S: Just the 4-2 whites and 2 blacks that might be remaining.

063 I: So you think it's--

064 S: That's contrived.

065 I: You don't--when you say contrived?

066 S: It's just pretty much--I don't really know.

067 I: You still think it's more likely to be white though?

068 S: I think it's more likely to be white, but--

069 I: Let me reach in--And, in fact, it's white. So now what do you think about which pocket I've been pulling out of?

070 S: [Points to A].

071 I: It's more likely to be pocket A.

072 S: Right. I would say pocket A.

<u>Lottery problem</u>

001 I: One last question. What would you be willing to pay for a lottery ticket to win a 10,000 dollar car?

002 S: Well, I've got a great volkswagon, so--

003 I: Yeah. Or 10,000 dollars cash prize?

004 S: OK. I don't know. A dollar I guess would be--'cause that's just what, uh, lottery tickets cost. That's just what I think of them.

005 I: Suppose in addition I told you that there were only going to be 100 tickets sold?

006 S: Only 100 tickets? Um, pro--divide the, uh--it's a 10,000?

007 I: Uh-huh.

008 S: By 100. And so that would be--it would be a 10,000--it would be a 100 dollars a ticket, I guess.

009 I: Uh-huh. Would you be willing to pay 100 dollars for--

010 S: I wouldn't be willing to pay. I mean, I guess that's similar to, uh, the house raffleing that's been going on,

011 I: Uh-huh.

012 S: where you pay 100 dollars for a 60,000 dollar house, or a chance at it. Uh--

013 I: You do think a 100 dollars would be a fair price to pay?

014 S: I--oh sure, if you wanted to--if you were into it--if you were into, you know, taking chances, uh, or if you, you know, I mean--

015 I: Would it be fair from both the perspective of the buyer and the seller of the ticket?

016 S: Um, well, it wouldn't--it would depend, um--No, I guess it wouldn't be fair to the person who was giving, you know, selling the ticket, because there'd be no profit in, um, in having the raffle.

017 I: Would it be unreasonable if someone were willing to pay 500 dollars for the ticket?

018 S: Um, for a chance at 10,000 and with the odds of 1 in 100?

019 I: Uh-huh.

020 S: Probably not. I don't think so.

021 I: Suppose you and a friend each bought a ticket, and you were present along with the other 98 ticket holders at the drawing. And just before the drawing you noticed a very tall person. They just caught your eye because of their height. Would you be more surprized in that tall person won the lottery or if you won it, or if your friend won it?

022 S: I'd probably--actually I'd probably be more surprized if I won it.

023 I: And why is that?

024 S: 'Cause, uh--I think I've won only one thing in my life, so--

025 I: Uh-huh.

026 S: And, and plus--well, that's also partly because I don't, um--I don't engage in that type of stuff, anyway. But, uh, I'd, you know, just thinking of it happening. I mean, if in any real sense it were to happen, it would probably be, probably pretty weird that it would happen.

027 I: You think you would be less likely to win it than, say, your friend or the tall person?

028 S: Probably less likely. I mean, as far as the numbers and stuff go, it would probably--if I was to--it would be the same, you know, for everyone involved.

APPENDIX C. INTERVIEW 2 TRANSCRIPT

Subject: S11
Interviewer: C. Konold
Date: May, 1982

Cab problem

001 I: OK. This is the first question and maybe you could read that aloud.

002 S: OK. A cab was involved in a hit-and-run accident at night. Two cab companies, the green and the blue, operate in the city. You are given the following data: 85% of the cabs in the city are green and 15% are blue. A witness identified the cab as a blue cab. The court tested his ability to identify cabs--wait a minute. I'm starting to forget information here already.

003 I: MmMm

004 S: --um, a hit and run--2 cabs--2 companies actually. Alright. More green as opposed to blue cabs. And now someone's identified it as a blue cab. The court tested his ability to identify cabs under the appropriate visibility conditions. When presented with a sample of cabs, half of which were blue and half of which were green, the witness made correct identifications in 80% of the cases and errored in 20% of the cases. What is the probability that the cab involved in the accident is blue rather than green? Whew!

005 I: OK

006 S: Let's see. Am I looking for a number as opposed to like--am I looking to say it's 80% probability that it's was a gree--er, blue rather than green? Is that what I'm--

007 I: What's the other option? How else would you prefer to give that?

008 S: "Sure it could have been a blue car."--No. Just that it would have been um--a strong--it was more likely as opposed to less likely--kind of like this fit in. More positive as opposed to a definite number positive.

009 I: MmMm. OK. Well why don't you give me the second of those first. You know, just your feeling about more or less, and then maybe you could try and give it a value.

010 S: OK. Sure. Um, first figure out what's the important data. And uh, this is a very tough one to me. Especially at 11 o'clock in the morning. So let's see. This may be important I guess.

011 I: The information that

012 S: That there's more green as opposed to blue. I'm not sure yet. Uh, and he was right 80% of the time in choosing a 50-50 split, which is important--Um, blue rather than green. Gee. Um, I'm trying to figure--just trying to sort it out. It seems as though in a sample where it's 50-50 and he's right very often, if in the real life situation there's more of the green, it seems as though the sample in the real life --would make this situation a little different, you know--His ability to predict would be changed, but I'm not sure if that's right. Umm, I think in making a decision, I would just take his ability to make correct identifications as an important piece of evidence.

013 I: MmMm

014 S: Umm, and I would say that--and if he said that--he identifies it as a blue cab, then the cab would be probably pretty close to the same type of probability as his ability to predict under a, kind of like a neutral type situation. So that I would say he's probably right. Er, I guess I would be strongly in favor of his ability to do it--

015 I: MmMm

016 S: to predict. And the probability would be somewhat close to the 80 to 20, I think.

017 I: MmMm

018 S: So that this piece of information might--it might affect it but I'm choosing not to use it.

019 I: MmMm. OK. So

020 S: That's a hard one though.

021 I: So you're feeling that it's--you're basing your estimate mostly on --entirely on the witnessess' ability to identify cabs under the test conditions.

022 S: Yeah, uh-huh.

023 I: So you feel he's probably right and the probability is around the 80%?

024 S: Probably around the 80%; yeah.

025 I: When you say 'around', are you--

026 S: Well, it's weird 'cause there are so few cabs out--blue cabs out in the real world that that may underscore the probability. It may alter it.

027 I: MmMm.

028 S: And I'm not sure how it would. Well, I would imagine that it would--it would reduce, you know, the probability that he was right,

029 I: MmMm

030 S: that it was a blue cab. Umm, just because there's fewer of them out there. And when there's a situation when he's got 50-50 and uh, he only makes a 20% error, but, I mean, if he had just been randomly guessing, he would've made 50% error.

031 I: MmMm.

032 S: And, you know, there's some difference between 20 and 50, but it's not--oh, it's significant but it's not uh, it's not as though this was like 99% right.

033 I: MmMm.

034 S: So when you change the, you know, the situation--the sample--then he may, er, this error could go up quite a but.

035 I: MmMm. So if anything, when you say 'about', you're thinking that it could be lower than 80%?

036 S: Yeah, I would er, it's weird. I'm not sure how this would in--I think in the direction of lowering the probability.

037 I: MmMm.

038 S: But I'm not sure if it really does. But it may. And so that I would say that it would probably go down to 70 or 65 or something like that.

039 I: MmMm.

040 S: Just because there's so few blue in the real world.

041 I: OK. Suppose the information up here in (i) were reversed and it said that 85% of the cabs in the city were blue and 15% were green?

042 S: OK. Umm, yeah, I think it would just be the reverse--that he would--The probability that a blue cab was involved would be higher than the 80, er, his ability to predict--so that'd be 80--maybe 85

043 I: MmMm.

044 S: or 90. Just because the probability that it was a blue car even irrespective of his ability, is higher.

045 I: Uh-huh.

046 S: Oh, that's an interesting way of looking at it.

047 I: What's that?

048 S: Well, um just omit 2 um, or double (i)--

049 I: Uh-huh.

050 S: that data. And just a chance that it was a green cab, if we come back to, eh, 85% are green and 15 are blue, just the chance that it was a green cab or a blue, actually a blue, cab on the spot was, er, is really, really low.

051 I: MmMm.

052 S: But it seems as though--I guess what I'm say er, I guess the way I'm thinking is that if the witness identifies it as a blue, and his ability is 80%, then chances are he was right that it was a blue cab, but without a witness it was probably a green cab.

053 I: MmMm.

054 S: 'cause there's so many green cabs.

055 I: MmMm.

056 S: But to jump back to when you said, well "is it important"--that if we jump back to er, "if 85% were blue, do you think"

057 I: MmMm.

058 S: Um, in that situation I think it would just be the umm, the probability that it was a blue cab would be really high. And if he's that good at predicting, umm, or, not predicting, but identifying which color the cab is, then it would kind of--the probability of it being a blue cab would be probably even better than 85%.

059 I: MmMm.

060 S: Maybe.

061 I: What do you--Why did you sort of, er--you were concerned suddenly when you said that.

062 S: Well, 'cause there's only, er, 85% of the cabs that are out there

063 I: MmMm.

064 S: Umm, are, well anyway, in the other--reverse situation

064 I: Yeah, in the reversed.

065 S: uh, they're blue. So, it seems weird to say that the probability that on that particular scene, there was a higher probability that it was a blue cab above the percentage of cabs.

066 I: Than the 85?

067 S: Yeah. But if this guy's good at predicting, then it's kind of like a--your're talking about his ability to predict blue cabs in that situation as opposed to predicting, you know, that it's red.

068 I: MmMm. Let me ask you this. Are you more sure when it's reversed like this--85% are blue--are you more sure that the probability is above 80% that it was blue than you are that the probability is below 80% when it's the way it is here?

069 S: Let's see. Above

070 I: Do you know what I'm--

071 S: in the reverse? Above 85?

072 I: When I gave you the first one you said 80--'maybe 80', and, er, you had some belief, but it seemed like a small belief, that it could be lower than that because of this. Reversed,

it seems to me like you're--you said 80 but you're even--you seem somewhat more sure that it will be even above 80%. I'm wondering if I was just picking up on that or--

073 S: Umm, let's see. I'm gonna go back to the one where, er, it's the normal situation. 85. It seems--well, I think it seems as though it would be the--it's the same. If 80 is like the midpoint of where they'll vary, you know, the situation as it is now, the probability will be, umm, let me see--Yeah. It seems as though the 15%--the 85 and 15% in the way it is now, in this normal situation, doesn't seem to, er, I don't take that into a strong consideration, this piece of information. And I go pretty much as if this is the sample that I'm talking about.

074 I: The 15?

075 S: The 15, the blue: Just his ability to predict.

076 I: Uh-huh.

077 S: And so I say it's kind of closer to 80%, the probability

078 I: MmMm.

079 S: in that situation. And when you reverse it, um, having 85% of the cabs green er, blue, yeah, it seems as though it influences it more in a positive direction.

080 I: MmMm.

081 S: But it shouldn't.

082 I: It shouldn't?

083 S: No, it should be the same. It's influence should be the same in either way.

084 I: Maybe you could tell me what you're thinking now as far as whether or not this information in (i) is relevant and should be taken into account in coming up with a value for the probability.

085 S: MmMm. Um, I think I--the way I explained it before that, um, in the real-life situation there's, um, a higher number of these, you know, there's a higher proportion of green as opposed to blue cabs

086 I: MmMm.

087 S: in the real world. And their chance of coming up is 50-50, in this one situation and, er, or in this situation. And then this guy's ability to predict is higher than chance, but it's not--it's somewhat higher

088 I: MmMm.

089 S: Um, but he still has quite a large error. Umm, so, it would--it doesn't seem that uh--

090 I: Just your best, er, what's your gut feelings about it--your

091 S: Well, it could influence it just because it's not--it's not the same as the 50-50 in which he's using as

092 I: MmMm.

093 S: in the sample. Um, so it's gonna adjust, er, it's a different type of situation under which he would be making identifications. So that would influence his ability to predict, er--well, not predict, but identify. But if this is a neutral situation and this 80 to 20 is just, um, a strong indication of just his ability to remember and to pick colors out, um, and that it's, uh, it's a meaning, er, you know, it's a meaningful number that is important to other situations other than the test situation. Then even in a situation where there's 85%, uh, green in the sample, he'll still be able to predict at an 80 to 20%.

094 I: MmMm.

095 S: So I think that if this means that he, er, it's not--it's, uh, meaningful to other than the sample, then this information is important. But if this doesn't mean that--

096 I: If this means that he--if this means that he can, er, that he will be this accurate in other situations, then (i) is or isn't important did you say?

097 S: (i) is not important.

098 I: OK.

099 S: Because--yeah. Because--

100 I: One last question. In giving the answer are you assuming--Are you making the assumption that the witness knew the relative number of cabs in the city?

101 S: No.

Bone-2 problem

001 I: You probably remember this from last time.

002 S: Oh, not that guy.

003 I: Yeah. Last time we did quite a bit of rolling with this bone.

004 S: MmMm.

005 I: And the first question I asked you about the bone was, if you roll it which side you thought most likely would land upright.

006 S: MmMm.

007 I: Do you remember what you gave as an answer?

008 S: No.

009 I: I'm going to ask you the same question again.

010 S: MmMm.

011 I: I'm not here to test your memory, but I want to change one thing. Last time you were rolling out of your hands. This time I want to ask the question what you'll get rolling out of the cup. Some people felt that rolling it out of the hand might make a difference.

012 S: MmMm.

013 I: And I don't think it did, but I don't want that to be an important question. Umm, and in answering that question--which side is most likely--I'm going to offer you any one of the following 6 pieces of information. Um, I'll give you a measure of the surface area of each side. Or the results of 100 rolls made by 16 people. That's a total of 100 rolls,

014 S: MmMm.

015 I: but 16 people contributed some part of that.

016 S: Allright.

017 I: The results I got rolling it 1000 times out of the cup. Uh, a drawing of the bone showing the center of gravity. Uh, the bone to look at. Or the results on your last 10 rolls.

018 S: Um, it's between 2 and 3.

019 I: Between--OK.

020 S: Um, Um, I'll go for piece of information number 3.

021 I: Mm, OK. And why do you take that?

022 S: Well beacuse it uh--1000 is a good number--No.

023 I: MmMm.

024 S: It's a lot of rolls. A lot of trials.

025 I: MmMm.

026 S: And um, it would seem that the natural er, the natural--the natural uh, whatever--lie or rolls that will come up--the most probably roll will probably show. The second prob--er, most probably roll will show--It seems like a good enough number--a high enough number to show the, uh, you know, the way the bone will roll.

027 I: MmMm.

028 S: You know, it's probabilities of the faces coming up. In number 2, I thought that because it was 16 people doing it, it would take any bias or whatever out of it. But I think it's probably pretty limited bias in rolling a bone anyway.

029 I: MmMm.

030 S: If it was like an attitude study or something, then I would probably go for the 16 as opposed to the 1.

031 I: And this is the second er,

032 S: Yeah.

033 I: If you had a second piece of information, you'd take number 2?

034 S: Uh-huh. Yeah, probably.

035 I: Is there any other information in there you think would be helpful?--or some that wouldn't be helpful at all?

036 S: Well, um, every, er, this would be pretty insignificant information.

037 I: The results on your last 10?

038 S: Yeah. But, because there's no--the 10 rolls isn't that--it's not that many,

039 I: MmMm.

040 S: to, um, to put trust in. Umm, and I remember looking at the bone and that didn't help too much. Heh-heh.

041 I: MmMm.

042 S: Uhh, I'll either, er, maybe the center of gravity would help or--yeah--oh, somewhat. I'm not an engineer so that wouldn't help me that much.

043 I: MmMm.

044 S: And the other one is related somewhat to this one. These seem to be the 2 more important pieces of information, and number 3 the most significant.

045 I: OK. Well, I'll give you both of those.

046 S: Oh.

047 I: Huh?

048 S: They're both generally pretty close to the same thing.

049 I: MmMm. So what would you say then, as to which side is most likely to land upright if you roll it?

050 S: B, C and D. Heh-heh. No. Oh, 1 side? It would just probably go, er, in either case just D will come up more often than the B and than C.

051 I: MmMm.

052 S: But this one never having come up. That's interesting.

053 I: So D is the most probable.

054 S: Yeah.

055 I: What's the probability of D coming up if you roll it?

056 S: Let's see. Almost 4 out of 10, or 40% or whatever.

057 I: MmMm. Now how are you getting 40%?

058 S: Just rounding up on this case.

059 I: MmMm.

060 S: And that, er, it might be worth going to--to 35, or cutting it right in the middle--$37\frac{1}{2}$ or something.

061 I: Uh-huh. But in that neighborhood anyway?

062 S: Yeah. Around 35 to 40%.

063 I: OK. Well, like last time I'm going to ask you to roll it 10 times--the bone.

064 S: OK.

065 I: But first to predict what you'll get--how many of each side that you'll get. And then to write it up there and keep track. So maybe you could do that.

066 S: So what am I going to do here? Write the numbers of A, B--

067 I: Well, no. You're going to roll it 10 times.

068 S: Oh.

069 I: And what I'm asking you to do is to predict beforehand what you'll get--how many of each side you'll get in 10 rolls.

070 S: Oh, OK. Hmm

071 I: OK. You're looking mostly at the results of 100 rolls there. Is that--

072 S: Oh, yeah.

073 I: Oh, you didn't know you were doing that?

074 S: I was just looking at it 'cause it's on my right.

075 I: Mmm.

076 S: Focusing right first, I guess. Hmm, you know it could influence the data. Let's see.

077 I: What are you trying to do in looking at that?

078 S: Umm, just assign a er, just reduce it to 10.

079 I: MmMm.

080 S: And assign a value for each one of these.

081 I: MmMm.

082 S: It feels pretty comfortable to go with these [B and D] as 3 and 3--3 and 3. And then this one [C] for 2. But it seems difficult to assign a 1 and a 1 for, er, 1 for A and 1 for E.

083 I: MmMm.

084 S: Just because they have such a low--probability of coming up. I will anyway--see what happens.

085 I: I don't think that attaches [referring to lid of felt pen].

086 S: It doesn't. OK. Write big? Small?

087 I: Large enough so it will er,

088 S: Is that too big?

089 I: That's fine.

090 S: [writes A-1, B-3, C-2, D-3, E-1]. Is that right? 2, 4 no--Yup, that's right.

091 I: So you've opted to give one to A and E

092 S: Sure.

093 I: even though you didn't feel quite--

094 S: Right.

095 I: How come you decided finally to do it?

096 S: Um, well I don't mind if I'm only like 90% or 80% right so, I figure one of these is gonna come up--Just spread the wealth and see which one does. Heh-heh. You can't go wrong.

097 I: OK. Well--

098 S: Oh, I guess the one that doesn't come up will go.

099 I: Go ahead and roll it.

100 S: [Rolls C] This is the A, right?

101 I: That's C.

102 S: Oh, darn-C.

103 I: You wanted your A right off?

104 S: Yeah. OK. So we've got 1 here.

105 I: Wouldn't you be concerned if you got your A right off. You might get 2 of them?

106 S: No. I don't know. Prob--yeah. Getting back to, uh, er, getting back to whether or not um, on each individual throw, it seems as though they have the same probability all over again,

107 I: MmMm.

108 S: of coming up. But then when you put all the throws together, it seems as though it has to work out to a pattern.

109 I: MmMm.

110 S: So that, umm, you would think if the A did come up first then, uhh, it could either fit into the pattern and it would only come up once, of it could come up again, and it could come again just because you're throwing it all over.

111 I: MmMm.

112 S: You're starting all over with the value, you know, the same probabilities. So, yeah, if it did come right, er, off right on--if it did come up right off, then I would think "oh-oh, could get 2 or 3," or whatever.

113 I: MmMm. So which, er, you mentioned a conflict there.

114 S: Yeah.

115 I: Having each roll--having everything be independent of the other one, versus fitting into the overall pattern. Which way do

you tend to think things happen?

116 S: Mm, on each throw it's all over again.

117 I: MmMm. OK.

118 S: Yeah, it's gonna happen again [rolls D] D. That's good.

119 I: That's good?

120 S: Yup. [Rolls B] Wow! It's working out like clockwork. Heh-heh. [Rolls D] D?

121 I: MmMm.

122 S: (Rolls D) D again? Oh no! So what, it--hmm. So it seems as though--it's weird.

123 I: What's weird?

124 S: Now I could get a million D's. So maybe that--maybe that's right. It's not the pattern--well, the pattern illustrates for each trial what the chances are to get the D or a C and a B.

125 I: MmMm.

126 S: But it seems as though in the overall thing you would just get about 35 to 37%.

127 I: MmMm.

128 S: Would be uh, would be D's.

129 I: So what's the conflict you're having right now?

130 S: It seems as though, er, when you have--when you get to the low--well, we're doing 10 here.

131 I: MmMm.

132 S: So when you get to the real low numbers of trials, and the probabilities are kind of like, er, they get away from the whole percentage points. And they get off into 37 or 35--well, 35 is fairly whole but

133 I: MmMm.

134 S: um, you can't have a half a roll or three quarters of a roll,

so that's going to influence the probabilities for the other values.

135 I: MmMm.

136 S: So that if this one comes up 4 times, you know, it's gonna influence how these others are gonna come.

137 I: MmMm.

138 S: I can't, er, I'm still not sure of the conflict between the individual trials and the pattern for the whole.

139 I: MmMm. You still feel the conflict?

140 S: Yeah. [Rolls B] Hmm, what kind of bone is this, again?

141 I: It's from a deer.

142 S: A deer. [Rolls B] B again. [Rolls D]

143 I: It sticks in there sometimes.

144 S: [Rolls D]

145 I: It's a D.

146 S: A D, oh-oh.

147 I: That's-that's, er, that's 8 rolls. What's your best guess as to what you'll get on the next 2?

148 S: Um, one of these two guys [A or E]. No, probably either a D or a B

149 I: MmMm.

150 S: on one of 'em and a C on the other.

151 I: MmMm.

152 S: And these guys [A and E], well, I should have left them out of it.

153 I: So you expect a D or a B and a C.

154 S: Yeah. And so I guess I'm going with the idea that--that they're not going to fit into the pattern but, well yeah;

It's just the individual trials that are gonna come up now.

155 I: MmMm. Do you think a C is more likely than a B?

156 S: Um, no. Probably B. Umm, they're pretty close to the B--the same chance of coming up.

157 I: MmMm.

158 S: But B has come up a lot more than C has so probably B has a better chance of coming up.

159 I: Based on what you've--right there, that it's come up more?

160 S: Yeah. Just based on, yeah.

161 I: OK.

162 S: 10 trials isn't that many. [Rolls B] Wow, B. Holy cow.

163 I: Now where do you put your money on this last roll?

164 S: The last roll? Let's look at it. Umm. Boy.

165 I: I mean if you were going to bet money on just what you get on this last roll?

166 S: Probably D.

167 I: MmMm.

168 S: Yeah.

169 I: Based on?

170 S: This and this [1000 data and last 9 rolls].

171 I: MmMm.

172 S: This is becoming, um, a factor I guess [last 9 rolls]. 'Cause I'm seeing how it's rolling now, and it seems to be influencing it.

173 I: MmMm.

174 S: [Rolls B] B would come up. Heh-heh.

175 I: So you would lose your shirt.

176 S: OK.

177 I: So what are you, uh, concluding?

178 S: Don't bet on A and E.

179 I: If you had to predict again for 10 more rolls, what would you predict?

180 S: Probably something like 4, 4 [D and B] and then 2 in here [C].

181 I: MmMm. And that's based on what you got there?

182 S: Yeah, I's try to base it on that. Although, I probably still want to go, actually, with a 3 here [B]. Beacuse in a thousand rolls, 279 came up.

183 I: MmMm.

184 S: Yeah, I'd probably still go with 3 here.

185 I: OK.

186 S: And a 2 [C] and a 4 [D]. Maybe--let me give one of these peripheral guys another break and give 'um a one.

187 I: OK. Let me put this aside here.

<u>Painted-die problem</u>

001 I: I've got a 6-sided die here. Suppose I told you that there was a possibility that it was loaded--that it had been altered in such a way so that such a one

002 S: Right.

003 I: side was slightly more likely than the others to land upright. Could you determine whether or not it way loaded?

004 S: I would roll it maybe. Yeah, it seems as though rolling it--rolling it in water, you know, it wouldn't--

005 I: Rolling it in water?

006 S: Yeah. Because if it was loaded and there was a heavy side or something,

007 I: MmMm.

008 S: maybe if you spun it around it would kind of--the heavy side down or something.

009 I: MmMm. OK. Umm, could you tell just by rolling it on that table, or something?

010 S: Yeah. It should come up more often on the, you know, the way you've loaded it.

011 I: MmMm. How many times would you roll, do you think, before you could conclude whether it was loaded or not?

012 S: Let's see. I was thinking of the 10 here [referring to the bone] that I rolled. I'm not real sure which sides are gonna come up. And that [the bone], in a sense, er--it's not loaded but there is a tendency of a certain spot er, face to come up.

013 I: MmMm.

014 S: Umm, so, probably like 20 or 30 or something like that.

015 I: MmMm.

016 S: You know--

017 I: Suppose I had rolled it 24 times and these were the results I got. You know, here are the number of times that each of the sides came up. What would you conclude looking at that? There's 24 rolls.

018 S: Hmm. It's weird how the 2 and 4 don't come up that often.

019 I: MmMm.

020 S: If it's a normal die, it has 6 faces. Umm, it should come up about 3 times for each, er, 4 times each; 4 times 6 is 24. Yeah.

021 I: MmMm.

022 S: So it should come up each side about 4 times if it was--let's see--Yeah. I guess like--I guess I would want to say that it might be loaded to the, you know--

023 I: MmMm.

024 S: to have the 3 come up, but it's not done in such a way that it's uh, it's not for sure the 3 is going to come up.

025 I: MmMm.

026 S: Although if it was 240 um, then we'd put zeros behind here. That'd be pretty significant. But, we don't have that information.

027 I: I see. So if this 240 and then I had 50, 20, 80--

028 S: Yeah, then it would be significant, but I don't think

029 I: MmMm.

030 S: I can say that, this being true.

031 I: MmMm. OK. In fact, as far as I know, this die isn't loaded.

032 S: OK.

033 I: Suppose I took this and I painted 5 of the surfaces black and left one surface white

034 S: MmMm.

035 I: and I put it in this cup. And rolled it 6 times. Do you think I'd be more likely in those 6 rolls to get 6 black surfaces or to get 5 black and 6 white?

036 S: Let's see. Probably, well, it's all centering on that conflict whether or not in the individual trials you start off with a fresh chance of having 6 er, have the 5 black come up as opposed to the one white.

036 I: MmMm.

037 S: But then at the end it seems as though after doing the whole, you know, 6 times, you should get 1 black--or 1 white.

038 I: Uh-huh.

039 S: I would say you have a higher chance of getting the six black.

040 I: MmMm.

041 S: I think I'll go with the idea that, um, on each trial you're starting off with the second--you know--a new set of probabilities--or the same, but starting fresh from each trial.

042 I: MmMm. OK. And that's why you say '6 black' because each trial you--

043 S: You'd get a better chance to get black.

044 I: you expect to get black. Umm, if I rolled it 60 times, how many whites would you predict I'd get in 60 rolls?

045 S: Umm, then, I'd want to just do it easy and say that because you've got a 5 to 1 thing

046 I: MmMm.

047 S: Um, 10 of 'em, or close to 10, would be white.

048 I: MmMm. But in 6 you predict I'll get--your best guess is no whites.

049 S: No whites; Yeah.

050 I: I don't--obviously I didn't paint the die like that, but I've got some familiar entities here from last time. I've got 5 black stones and 1 white stone. Suppose I put these in the cup and shook it up real well and reached in, not looking, drew one out, wrote down the color, put it back in, shook it up again and kept drawing like that.

051 S: MmMm.

052 I: Would that be the same as rolling the die that I described that was painted?

053 S: Yeah, I think so.

054 I: You think so?

055 S: Yeah. You've got six face or colors--

056 I: MmMm.

057 S: 6 whatever. Sides--if you want to call them sides.

058 I: MmMm.

059 S: They're stones in this particular case, but there's 6. And you're drawing one of 'em from the six.

060 I: MmMm. Umm, you know, if I rolled, say, the die on the table and I kept track of what I got, and then I drew these stones out of here and kept track and I showed you the results, would you be able to tell which was from the die and which I'd gotten from drawing these stones?

061 S: I don't think so.

062 I: Well, I'm going to, in fact, draw 6 times from this and ask you to predict what I'm going to get. First of all, what do you guess overall that I'm going to get from 6 draws.

063 S: 6 black.

064 I: 6 black: OK. I've got the first one.

065 S: Black.

066 I: And what's the probability?

067 S: Umm, probability? Uh, and well, 5 in 6.

068 I: 5-

069 S: 5 out of 6; whatever that would be.

070 I: What do you mean 'whatever that would be'?

071 S: Umm, just whatever percentage.

072 I: Oh, whatever percentage: OK. 5 out of 6 is fine.

073 S: 5 out of 6.

074 I: OK. There it is--it's black. OK.

075 S: I keep going with black.

076 I: What't the probability?

077 S: 5 in 6.

078 I: OK. Well I'm going to made you do this. There's two [showing the 2nd black] OK.

079 S: It's gonna be black.

080 I: What's the probability?

081 S: 5 in 6.

082 I: [Showed black] That's 3.

083 S: Same thing--5 in 6 black.

084 I: [Showed black] Four. Are you uncomfortable doing this?

085 S: Ya, cause this guy's gonna come up white, Heh-he, ha-ha.

086 I: OK. Here's another one.

087 S: It's still going to be black, and it's gonna be 5 out of 6.

088 I: OK. Is that 4 or 5?

089 S: I think it's

090 I: 5

091 S: 5, yeah.

092 I: OK. One more. OK.

093 S: It's gonna be black and it's gonna be 5 out of 6, whatever.

094 I: Ha-ha [showed black]

095 S: It's er, it's black.

096 I: Um, now if I did that another 6 times,

097 S: MmMm.

098 I: Would you expect more likely that I get 6 black, or 5 black and 1 white?

099 S: Um, 6 black. Yeah, I think so.

100 I: So you're always going to go with 6 blacks?

101 S: Yeah, but somewhere along the line, the white is gonna come in.

102 I: MmMm.

103 S: Every once in a while, but I'm not sure, er, I'm not sure how it figures into the situation though.

104 I: MmMm.

105 S: Whether or not the--the white uh--it seems when you just--
when you get a large enough samples or trials, then you have
the influence of the white. But under the, uh, the smaller
samples, um, it seems as though--in this case, the black,
the one with the highest probability of coming up, er,
being chose, becomes more important--

106 I: MmMm.

107 S: you know, significant. I don't know.

108 I: OK. Umm, you agreed that instead of rolling the die that I
described, that had 5 black surfaces and 1 white, that it
would be comparable--I'd get the same results if I put
these stones in the cup.

109 S: MmMm.

110 I: So, therefore, I've somehow made a model of this die with
these stones.

111 S: Right.

112 I: Would there be a similar way that I could model the bone so
that instead of rolling the bone I could pick something out
of, like an urn, and get

113 S: Hmm.

114 I: the same kind of results that I do from rolling the bone?

115 S: Well, probably I wouldn't, er, I would say no beacuse uh,
prob--these are all the same you know. These sides are
all the same.

116 I: MmMm.

117 S: But the faces of this bone aren't the same.

118 I: MmMm.

119 S: Or their probability of being chosen isn't the same.

120 I: MmMm.

121 S: Unless you could alter the probability of like, er, like of
a particular type of stone to be chosen.

122 I: MmMm.

123 S: And I would say no.

124 I: Any idea of how to alter the probability of a particular stone?

125 S: Hmm. Well, let's see. In some way maybe you could physically change it to make it--I don't know. Naw, I can't think of a way to do it.

126 I: Let me suggest a couple of things here. I could, for example, label -- put little labels on the stones,

127 S: Right.

128 I: one for each side. Oops. Now I've got 6 over here

129 S: MmMm.

130 I: And if I put these in there and shook and drew, would I get results like the bone?

131 S: No.

132 I: That's sort of what you were talking about.

133 S: Right.

134 I: Well, is there some container that I could fill with some number of lettered stones that would give me results like the bone?

135 S: Oh, I see. Yeah. Maybe like, um, the opening: Maybe the size of the opening um--Like in this case, A and E maybe will just exactly fit the size of the opening of the container

136 I: MmMm.

137 S: so that it would take a near perfect throw and, uh, those particular stones would have to be at the front of the line leaving the urn.

138 I: MmMm.

139 S: And, uh, having the other faces like--like D and B, er, well B would be probably the smallest and maybe easiest to get out, and then D would be a little bit more difficult or harder to get out, and C would be even more difficult and then A

and E would just have to be perfect.

140 I: I see. So we'd put some covering on this that had a er, the hole size.

141 S: Yeah, something like that.

142 I: Um, and then these would be different sizes.

143 S: Yeah, right. 'Cause the faces are different,

144 I: MmMm.

145 S: so different that, uh, it affects their chance of coming up. So you'd want to do that.

146 I: So do you think I could engineer that such that I could then

147 S: Sure.

148 I: do that so that you couldn't, er, so that I'd get results just like rolling the bone?

149 S: Sure. Farm out the project to the engineering department.

150 I: Let me suggest another alternative to you that somebody--

151 S: OK.

152 I: If I didn't suggest it, I wouldn't get to show you all my beads.

153 S: OK.

154 I: Now suppose, um, suppose we took the bone to a statistician and however they decide it he decided that the probabilities, um, were very close to these values.

155 S. MmMm.

156 I: That A was about 5 out of a hundred or 5%; b 29 out of 100, 29%; C 24; D 37; and E 5 and F zero. And so we took a big can and we took five of these A's and put them in here. And took 29 B's and put

157 S: MmMm.

158 I: them in there, and 24 C's, 37 D's, and 5 E's--I'm just taking those values. And then we shook and drew from that. Do you think that would give me results comparable to rolling the bone?

159 S: Yeah, that would be pretty close. But I would feel uncomfortable about not having F represented.

160 I: MmMm.

161 S: But yeah, that would, er, that's--that would seem as though it's a fairly--a good way to do it.

162 I: MmMm. OK. Um, and if I did that again under the table and I rolled the bone like 100 times and kept track,

163 S: MmMm.

164 I: And I drew 100 times from this and I showed you the results, do you think there'd be any way for you to tell which results I got from the bone and which from drawing from this urn?

165 S: Umm, yeah. If F did come up once, that would give it away.

166 I: Yeah.

167 S: You know, if F didn't come up, probably not.

168 I: MmMm. In those 100 trials, you know, that I got from the bone and from drawing from the urn, do you think with one of those I'd be more likely to get no E's? Like rolling--like would I be more likely to get no E's rolling the bone, or be more likely to get no E's drawing from the urn?

169 S: I think you'd get more--Let's see.

170 I: Or equal?

171 S: Yeah. It would seem as though in this situation [container] you would, uh, you'd have fewer E's.

172 I: I'd get fewer E's in 100 trials with the urn?

173 S: Yeah, with the urn. Yeah.

174 I: Why?

175 S: 'Cause there's so many--gees--it's kind of like, uh, well, I mean there's so many in this group as opposed to the bone. I don't know, something about the bone and there being so many, many, many and many of these guys [D's] coming up.

176 I: Uh-huh.

177 S: Uh, it's weird. It's like, um, if you could al--if you could alter when you have one of each and, um, the probability of them coming up is not related to the number.

178 I: Like the one you described first with the opening?

179 S: Yeah, Yeah. It's weird. But I don't know--it's strange, 'cause, uh, if you--if you reduced the population as you took them out, you know, then it would probably, you know, I mean it would obviously fall into the nor--the same pattern. But--

180 I: What do you mean 'if you reduced the--'

181 S: As you got a D, you took it out.

182 I: Oh.

183 S: Then you'ld give these guys a break.

184 I: Yeah?

185 S: Then they could come up.

186 I: Uh-huh.

187 S: But, it seems in this situation, you would get a lot of Ds and Bs coming up, and some Cs.

188 I: Do you think I'd get more Ds and Bs in here than I would get with rolling the bone?

189 S: It would seem as though you would.

190 I: So, you might be able to tell the difference then.

191 S: Um, yeah. I hadn't thought of it--Yeah. Yeah, beacuse it's like you're only going to take one of these, uh, stones out at a time.

192 I: Uh-huh.

193 S: So, you take the D, or whatever, the D or the B or--you take one out and put it back in. Seems as though you'll take a lot of these guys [B and Ds]. But the bone's the same thing, I guess. 'Cause when you roll it, um, these two faces come up very often.

194 I: Uh-huh.

195 S: [Looking at the bone and numbers] So I'm trying to decide--what I'm trying to decide is whether or not, just the numbers is a good way to represent the prob--uh--the chances or the probability that these guys will come up in the natural situation.

196 I: Uh-huh.

197 S: And, on one side I say it seems as though that they wouldn't, because--well, I don't know. Maybe they--I think it actually --it would, because it just represents the chance of them coming up, and since there's a lot of these guys here--but there's, uh, in the bone situation, there's just a real high chance--or just through natural rolling--that the D and C and B will come up the most often.

198 I: Uh-huh.

199 S: So that I think what it really is, it just--this is a way of representing the natural--the natural differences in the probability of the bone. So, actually, I think I'll change back around and say that if you did it both under the table, I wouldn't be able to tell the difference.

200 I: Uh-huh. OK. So if I did this 100 times--drew from the can with these, how many Es would you guess I'd get?

201 S: Um, 100 time? Four of five, something like that.

202 I: Uh-huh. And if I rolled the bone 100 times, how many would guess?

203 S: About four or five.

204 I: OK. Um, one person suggested another way to do it too. It's a slight modification of this--that I take the bone and roll it out of the cup, and I take that letter, whatever I got, an E, and I put it in there. And then I do that again, and I roll it, and I get a B, and I put a B in there.

205 S: Right.

206 I: And I just keep rolling it

207 S: Uh-huh.

208 I: a long, long time. Uh, would I reach a point when I could start drawing from the can full of the beads that I'd put in there in that manner,

209 S: Uh-huh.

210 I: and get results comparable to the bone?

211 S: Yeah. I think that would happen.

212 I: Uh-huh. Would that feel better than the--

213 S: No. I think it's the same,'cause this is the same thing, 'cause it's just done--you did all your tab--you rolled the bone and then you filled in the--

214 I: Uh-huh.

215 S: And this was a statistician's, or whatever, probabilities, right?

216 I: Uh-huh.

217 S: Oh. Oh, that's right. Oh, well then I would go for the bone rolling one, because that's more of the natural roll, of the--of the bone.

218 I: Uh-huh. How do you think a statistician would come up with the probabilities of the bone?

219 S: Um, hopefully he rolled the bone.

220 I: Hopefully, huh?

221 S: Yeah.

222 I: What do you mean 'hopefully'? That he didn't just invent them, you mean?

223 S: Yeah, or he didn't just make them up, you know--I guess.

224 I: If we took this bone to a statistician and we wanted them to figure out as precisely as they could the probabilities

of each side,

225 S: Uh-huh.

226 I: how do you think the statistician would go about doing that?

227 S: Um, well the way I would--I'd just roll it a lot of times to see which--which faces came up.

228 I: Uh-huh. OK.

229 S: But maybe he would take it to the engineering department and figure out the center of gravity, the--and etcetera, you know.

230 I: Uh-huh.

231 S: The properties of the bone that--

232 I: 'Maybe he would', you're not sure?

233 S: Well, it depends, you know, on how deep the guy is, you know.

234 I: Which would be--in your mind, which

235 S: Both.

236 I: is the best way?

237 S: Both. It would be a joint effort.

238 I: Uh-huh.

239 S: 'Cause, um, you'd get the--you'd roll the bone and get a rough idea of the probabilities, or even a fine-tuned idea of the probabilities, whatever they are, yeah, probabilities--and take it to have it analyzed to figure out if, structurally, um, you can understand why these--why these, um, you know--you assign these particular values to each face. And then through comparing both, just--

240 I: But I might want to modify what I got rolling it?

241 S: Yeah. Just--it's just kind of like added significance, or not significance--added, um, sureness, or whatever--belief in your percentages.

242 I: Suppose the statistician had rolled it 1000 times, and, you know, he was on a fixed budget, and he could either roll it 1000 more times, or take the thing over to the engineering department and have them look at it. Which do you think he'd be better off doing?

243 S: Um, I'd go with the corroborating evidence from the engineers.

244 I: Um, rather than just duplicating it?

245 S: Mm.

246 I: OK. That's all the questions I've got.