UNDERSTANDING STUDENTS' BELIEFS ABOUT PROBABILITY

CLIFFORD KONOLD

University of Massachusetts, Amherst

Probability is a particularly slippery concept. What makes it uniquely troublesome is the territory it stakes out. Through probability, we attempt to demarcate the amorphous state somewhere between the imagined extremes of 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, the conceptual landscape not only trips you up, but keeps you sliding once you're down.

According to Hacking (1973), prior to the 17th century such a middle ground between belief and knowledge did not exist, yet the term *probability* had already been around for some time. Hacking's thesis is that in the early to mid 17th century 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 that could only be testified to either by men of authority or through God-given signs (*opinio*). The word *probability* was associated with the latter class such that an opinion was probable if it had received the stamp of approval of some authority. This meaning was still around as late as the early 18th century as manifested in Daniel Defoe's 1724 novel *Roxana*, in which a woman describes her living arrangements thus:

> This was the first view I had of living comfortably indeed, and it was a very probable way, I must confess, seeing we had very good conveniences, six rooms on a floor, and three stories high.

A "probable" belief (or circumstance) was thus an "approved" one. But no amount of approval could ever make an approved belief a matter of demonstrable knowledge. Hacking argues that a bridge was created between belief and knowledge as a result of a change in, or enlargement of, the concept of evidence. 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 (like coin flips or die rolls) to produce stable frequencies over many repetitions. This dual usage was no accident, for frequency data were a special variety of this new type of evidence that permitted degrees of belief separate from opinion. Whether a belief was probable had now less to do with approval of authority than with the approval of data.

E. von Glasersfeld (ed.), Radical Constructivism in Mathematics Education, 139–156.

If we were to meet a 16th century ancestor, we would no doubt get bogged down trying to understand what we each meant when we said that something was probable. Unfortunately, we need not time travel to observe or engage in a confusing dialogue about probability: We need only to walk into a classroom where probability is 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" versus to "believe." Similarly, I will argue in this paper that understanding how students think about probability before and during instruction can facilitate communication between the student and teacher of probability.

As other contributors to this volume stress, information cannot be picked from the external environment like apples from a tree. What a student mentally carries away from a teaching episode is not information that is received directly, but that is constructed. The construction involves the student weaving *selected* and *interpreted* teacher outputs into an existing fabric of knowledge. What a student learns from a given classroom experience is both limited and, at the same time, made possible by what he or she already *knows*.

Of course, the role of prior knowledge in learning is only half the picture. If I believed that there were no mechanism by which the external environment could impose itself in some way upon a person's conceptions, I certainly would not be writing this paper, nor would I be concerned with any other social, collective activity. I would be condemned, or delivered, depending on one's point of view, to my own internally isolated world. The nature of this "accommodating" mechanism is discussed by Piaget, von Glasersfeld, and others, and has to do with the fact that some conceptions allow me to accomplish certain goals while other conceptions do not. The fact that I cannot pick up an apple until I have learned to accommodate my hand to certain of its properties prevents me from treating the apple in a totally arbitrary way if I indeed want to grasp it. Now, of course, I'm arguing that we in fact do acquire information from the environment much like we pick apples from trees. But I hope that in the process of contradicting myself, I have alerted the reader to what might be involved in the act of apple-picking – that it involves a reciprocal relation between the apple and the picker which is more complex than we ordinarily grant.

Constructivist educators are of the opinion that the student-teacher relationship is more complex than a casual analysis would lead us to believe, and that while apple-picking can continue well enough without critical analysis, teaching cannot. While this is a debatable point, my purpose here is not to defend that opinion. I state it here as an assumption and proceed instead to demonstrate how investigations into student understandings of probability can be of pedagogical value.

Although I don't want here to defend my point of view, it will serve my purposes to exemplify what I regard as a contrary one. The following quotation is from a chapter in Cherry (1968) and is part of an introduction to Carnap's logical-relations theory of probability:

> As with all mathematics, there is no need of definition of the basic concepts, but only of the formulation of rules for using them. We do not need to know what probabilities "are," but rather how to combine them. But, nevertheless, it helps most people to have intuitive notions of the basic concepts, though such notions can become a hindrance to the expert (p. 234).

This quotation is grounded in a particular philosophy of mathematics, a theory of mathematics education, and a view of mathematical expertise. The philosophy of mathematics is one that has dominated mathematics since the 1920s. According to it, mathematics does not address questions concerning the interpretation or meaning of mathematical objects, but focuses attention instead on the validity of operations on those objects. Mathematical expertise thus involves facility with abstractions, and not necessarily with being able to apply the mathematics to some real-world problem. This philosophy has encouraged the view that to engender mathematical thinking, the teacher ought to divorce the activity in the first place from real-world experience. While it may sometimes help to provide some grounding intuitions, these are viewed as temporary supports rather than as foundation stones. Furthermore, these intuitions seem to be regarded as notions that the teacher may choose to develop.

My assumption is that students have intuitions about probability and that they can't check these in at the classroom door. The success of the teacher depends in large part on how these notions are treated in relation to those the teacher would like the student to acquire. Additionally, I think it is a myth that mathematics, either as a discipline or in the mind of a mathematician, develops independently from concerns about objects and relations that are believed to have real-world referents. This was certainly not so in the case of the development of probability theory.

I. FORMAL INTERPRETATIONS 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 fairly 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 the statement, "Now our harmony has begun again" (Maistrov, 1974, p. 39).

What came into being at that time was a new tune; not so new were the lyrics – "How probable is it that ...." Since that time we have been plagued (some would argue) with a concept of probability that has at least two different, but related, aspects. On the one hand it is a degree of subjective belief in the truth of some proposition; on the other, it refers in a precise way to an objective property (frequency of occurrence) of certain types of events.

There are several schools of probability, each with a somewhat different interpretation of probability. The first interpretation to which students are typically exposed is the *classical interpretation* according to which the probability of an event is simply the ratio of the number of alternatives favorable to that event to the total number of equally-likely alternatives. Thus, the probability of rolling a 3 with a fair die is 1/6, since only one alternative out of a total of six is favorable to the outcome. Introductory textbooks often refer to the classical interpretation as "theoretical probability." This interpretation is obviously limited to trials with objects such as coins, dice, and spinners, that are composed of equally-likely alternatives. It is also logically flawed in that its definition of probability is circular: Probability is defined in terms of equally-likely alternatives, yet what can be meant by "equally-likely" other than "equally-probable?"

Perhaps the most pervasive objective theory of probability is the *frequentist interpretation*, which defines probability in terms of the limiting relative frequency of occurrence of an event in an infinite, or near infinite, number of trials. Thus the probability of rolling a 3 with a particular die is the ratio of trials resulting in a 3 to the total number of trials (actually the limit of this ratio as the number of trials approaches infinity). In introductory textbooks, the frequency interpretation is sometimes referred to as "empirical probability." While this interpretation can be applied to events that are composed of non-equally-likely alternatives, it is restricted to setups such as coin tossing and urn drawings with which "identical" trials can be repeated indefinitely. In addition, since probability is defined as frequency in the long run, it is meaningless from this perspective to talk about the probability of an event occurring on a particular trial. It should be pointed out that even though this interpretation is referred to as an objective theory, this does not mean that it is free from subjective considerations. Most importantly, this interpretation requires an observer who can count a series of events, and in order to accumulate a sum of occurrences, that observer must consider the events which are counted to be of the "same type." If each flip of a particular coin is considered unique in an essential way, and there are certainly grounds for doing so, then the frequency interpretation cannot be applied.

One advantage of various *subjectivist interpretations* of probability is that they can be applied to a wide range of phenomena since, according to this view, probability is a measure of belief in the truth of a proposition. According to a subjectivist view, different people could validly assign different values to the probability of rolling a 3 with a particular die. These values would presumably reflect different beliefs about the fairness of the die, the character of the person doing the rolling, the technique used in rolling, etc. However, in formalizing subjectivist interpretations, theorists have adopted various adjustment mechanisms (e.g., Bayes' theorem) that lead to the revision of initial probabilities given new information such as results of actual trials. Given enough data about the frequency of occurrence of 3 with a particular die, the various subjective probabilities held by different people would all begin to converge on the frequentists' limit.

The meaning of the probability value in a subjectivist interpretation can be thought of in several ways. One of the most common is to describe the value as a measure of a person's belief in what would constitute a fair bet. Thus, a person estimating the chance of rain at 70% would as quickly bet $7 to win $10 if it did rain as they would bet $3 to win $10 if it didn't rain. Another way to think about the meaning of the value is to consider as a collection all those events to which a person would assign the probability 70%. We could then look to see how many of these predicted events actually occurred. If only 50% of the events which had been assigned 70% probability actually occurred, then we would conclude that this person was, in general, overconfident in his or her predictions. If 90% of the events actually occurred, then the person would be underconfident. The nearer the percentage of predicted events came to 70%, the more expert we would regard the person in arriving at subjective probability judgments. In the lingo of decision theory such a person would be "well-calibrated." Thus, if we assume that a person is well-calibrated, when this person assigns the probability p% to events we should expect roughly p% of those events to occur.

Notice that the subjectivist theories are *normative* theories, specifying how "rational" people ought to formulate and alter their beliefs in the light of new information; they are not *descriptive* theories of how people actually formulate and alter subjective probabilities. I use the term "normative" in this chapter to refer to theories or beliefs that are held by those regarded as experts. When I speak of "normative theories of probability," I will be referring not to any particular theory of probability but collectively to all those theories that are considered by experts in the field as deserving serious attention. Thus, the term "normative" is not synonymous with "correct," but is very similar to the pre-modern meaning of probable, i.e., "approved."

These various interpretations are mentioned here primarily to provide a context for evaluating various beliefs about probability held by students. I want to proceed by suggesting that two elements of Hacking's account of the development of the modern conception of probability have particularly important implications for the teaching of probability. First, probability is viewed by Hacking as a concept that exists in a web of discourse and related concepts. Thus, the modern concepts of frequency of occurrence could not *emerge* until the related concept of evidence changed. Second, once a new, objectified notion of probability had appeared, it wasn't totally divorced from its prior, epistemological meaning. Hacking claims that concepts "play out their own lives in, as it were, a space of their own" (p. 15). Concepts are not freely created and revised at will. Rather, they are subject to restrictions that are inherent in the web of concept-relations in which they are embedded. The meaning of a word cannot be directly enforced; it emerges from and is supported by a tenuous consensus among those who use the word. Thus, despite efforts of advocates of both objective and subjective schools of thought to annex the concept of probability, it maintains even today its Janus face.

I want to argue that these two features also apply to communications between students and teachers regarding probability. The teacher has a particular concept of probability and intends to communicate, in some way, this concept to the students. Ideally, the teacher's view is similar to one of the interpretations summarized above. To the extent that the conceptual frameworks of the teacher and the student are compatible, a sense of mutual understanding is possible. Where those frameworks are incompatible, however, communication becomes problematic. The teacher cannot, by decree, enforce a normative view. The primary reason for this inability to simply transfer a concept from teacher to student is that students have no other option than to *interpret* what the teacher says or does in light of what they already understand of probability. As Fisher and Lipson (1985) have written:

> ... what we perceive is strongly influenced by our beliefs and expectations. The constructive nature of thought has the consequence that every observation is an interference and is dependent upon our existing knowledge of the world. In spite of this, we frequently become aware of differences between our expectations and observations. As a result we are more or less constantly engaged in assessing the "goodness-of-fit" between our mental models and our experience with the world around us (p. 50).

Not only does this quotation clearly express the idea that what we perceive and experience as existing in the external world is dependent on our prior knowledge, it also inadvertently demonstrates the very point. The word *interference* in the second sentence ought to read *inference*. Whether or not a reader notices and resolves this error depends on prior knowledge. Some readers, like the authors of the article, who are very familiar with the expressed idea, are likely to see what they expect to see and read the word as *inference*. Others read the word as *interference*, but knowing the point that the authors are making can fairly quickly replace it with the intended word. In this way they validated the claim made in the last part of the quote. Those totally unfamiliar with the authors' view of perception are more likely to read the word as *interference* and perhaps create an interpretation with which this word fits.

Long before their formal introduction to probability, students have dealt with countless situations involving uncertainty and have learned to use, in common discourse, words such as *probable, random, independent, lucky, chance, fair, unlikely*. They have an understanding that permits them to use these words in sentences that are comprehensible to other language users in everyday situations. It is into this web of meanings that students attempt to integrate and thus make sense of their classroom experience. Unfortunately, since much of formal probability theory does not fit well into the informal understandings that students already possess, misunderstandings, serious confusions, and other breakdowns in communication occur. If a teacher can come to see how students view probability, that teacher will be better able, emotionally as well as conceptually, to aid the development of a normative concept. With this goal in mind, I describe in the next section some of the non-normative beliefs and perceptions about probability that have been found to be commonly held.

II. STUDENT CONCEPTIONS

The view that statistics and probability are conceptually difficult topics for the untutored is relatively recent. Up through the 1960s it was generally assumed that when people reasoned in everyday situations about uncertainty they used methods similar to those a statistician would use, but less carefully (Peterson & Beach, 1967). Piaget and Inhelder (1951) claimed that by the age of 12 most children could reason probabilistically about a variety of randomizing devices and had developed sound statistical notions including an intuitive understanding of the Law of Large Numbers.

More recently, cognitive psychologists have found pervasive and persistent errors in peoples' reasoning under uncertainty. Contrary to the earlier assumptions, research indicates that people arrive at probabilistic judgments via considerations that are qualitatively different from the statisticians'. Daniel Kahneman and Amos Tversky (1973) have offered the most comprehensive account to date of why peoples' judgments often differ from accepted theory. They suggest that because of limited information-processing capabilities, people use various judgment heuristics that allow them to summarize large amounts of data and quickly arrive at decisions.

Although these heuristics usually produce adequate estimates, they are limited in the amount and type of information to which they are sensitive. This, in turn, leads to predictable judgment errors in some situations. For example, most people incorrectly believe that the ordered sequence MMMMMM of male and female births in a family is less likely than the ordered sequence MFFMMF. One possible explanation of this belief is that people confuse the latter sequence with the unordered outcome "3 Ms, 3Fs." If the precise order of births is disregarded, it is certainly the case that the outcome "3 Ms, 3Fs" is more likely than "6 Ms." (There are 20 different ways to arrange 3 Ms and 3 Fs and only 1 way to arrange 6 Ms.) Kahneman and Tversky (1972) suggest, however, that MFFMMF is judged as more likely through the application of what they term the "representativeness heuristic." According to this heuristic, the probability of a sample is estimated by noting the degree of similarity between the sample and parent population. Since the sequence MFFMMF is more similar to the population proportion of approximately half males and half females, and also appears to better reflect the random process underlying sex determination, it is judged as more likely.

The heuristics described by Kahneman and Tversky are useful in explaining why so many results of probability theory seem counter-intuitive and incorrect to the student. In fact, it has been demonstrated that even experts in probability can be led to the unconscious application of heuristics for situations that they know call for a probabilistic analysis (Tversky & Kahneman, 1971). Heuristics thus appear to be deeply held and, in many cases, automatically applied. They have been compared to optical illusions where, even though one may know better, the situation cannot be perceived veridically.

Another way to describe these heuristics is that they provide an account of how people with no formal training derive an estimate of the probability of an event. Hidden in the heuristic account is the assumption that regardless of whether one uses heuristics or the formal methods of probability theory, the individual perceives the goal as arriving at the probability of the event in question. While the derived probability value may be non-normative, the meaning of that probability is assumed to lie somewhere in the range of acceptable interpretation.

I have been studying a pattern of errors that are more fundamental, in some respects, than those resulting from the use of judgment heuristics. Rather then exploring how people *arrive* at probabilistic judgments, I have been interested in how people *interpret* a question about probability or a probability value. Results of interviews in which college students reasoned aloud about situations involving uncertainty indicate that for many students a question about probability is interpreted in a way that is, from a normative view, non-probabilistic (Konold, 1989).

According to this alternative interpretation, which I will refer to as the "outcome approach," the primary goal in situations involving uncertainty is not to arrive at a probability of occurrence but to successfully *predict* the outcome of a *single* trial. Given this objective, a question that explicitly asks for the probability of an outcome is interpreted as asking whether the outcome will, in fact, occur on the next trial. For example, asked to explain a weather forecaster's prediction of 70% chance of rain, many students respond that they take that to mean that it *will* rain. Asked what they would conclude if it did not rain, these same students hold that the forecaster's prediction would then have been wrong. They also will argue that a forecaster is performing sub-optimally when it rains on 70% of the days for which 70% chances were given. Probability values are evaluated in the outcome approach in terms of their proximity to the anchor values of 100%, 0%, and 50%, which have the respective meanings of "yes," "no," and "I don't know." Thus students reasoning according to the outcome approach will argue that 70% is sufficiently close to 100% to warrant the assertion, "It will rain tomorrow."

Students who show this preference for predicting individual as opposed to sample outcomes also tend to rely on causal as opposed to chance explanations of outcome occurrence and variability. For example, some students believe that the number in the estimate "70% chance of rain" refers to a measure of some factor that causes rain. Thus they believe that the 70% implies 70% humidity, or

70% cloud cover. While predictions are often based on inferences about causality, frequency information is often ignored. For example, I gave students an irregularly-shaped bone to roll and asked them which of the six surfaces was most likely to land upright. As mentioned above, many of the students appeared to interpret the question as a request to predict the outcome of a single trial and accordingly evaluated their predictions as being correct or incorrect after one roll of the bone. And, remarkably, after conducting several trials with the bone and being shown a summary of results of 1,000 trials, some students preferred to base their predictions on a visual inspection of the bone rather than on the available data.

In the remainder of this section I will present and discuss in some detail a few excerpts from conversations with students about probability. My purpose is to demonstrate how statements that otherwise would be regarded as contradictory or incomprehensible can be understood if one assumes that students are reasoning according to the outcome approach. More generally, I want to suggest that understanding how students are thinking about a topic – in this case probability – puts the teacher in the position of being able to initiate conversations and design curriculum that can facilitate the development of more adequate concepts of probability. Without this type of careful attention to student thinking, instruction in probability will, for many students, have no lasting impact.

2.1 Making Confusing Statements Understandable

The excerpt presented below demonstrates how statements that would otherwise seem disconnected and incomprehensible can be seen, with the aid of a model of student reasoning, as both comprehensible and connected. The conversation took place in an experimental setting in which the student had been asked what a 70% chance of rain meant, and is here responding to the follow-up question:

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?

The student first responded to this probe by concluding that the forecaster was "very accurate" but then began to question whether, in this instance, one could look beyond the prediction of an individual day in assessing accuracy:

S: 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 setup.

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:

S: 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
S: Yeah.
I: accurate? In terms of what?
S: I guess, as I said, he must have certain standards to go by when he picks a chance of rain.

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

I: Say the forecaster predicted 50% chance of rain on those ten days. What would have happened if the forecaster was really good? 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?
S: 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 –
S: 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.

These rather confusing statements become fairly comprehensible when the student is viewed as interpreting the 50% in a way that is consistent with the outcome approach. According to the outcome approach, 50% is the mid point of the yes/no decision continuum, and thus it means that anything can happen – "It may rain, or it may not rain. Who knows?" Under this interpretation the 50% carries no implication concerning the frequency of rain. As a result of considering the meaning of a 50% forecast, this student confirms her conclusion that assessing accuracy over days is inappropriate. She is apparently aware of the implications involved in assessing the accuracy of a forecast of 70% chance of rain by looking at the frequency of rain on days for which a 70% chance of rain is given. To be consistent with this, one would expect a 50% forecast for 10 days to be accurate if it rained on 5 of those days. But this interpretation of accuracy would contradict the use of 50% as an indicator of "anything-could-happen" since according to the latter interpretation, any imaginable result of a ten-day sample would seem equally consistent with a forecast of 50% rain. This reasoning makes her more certain that individual days is the appropriate unit of analysis, and accordingly, when asked to summarize what she believed, she responded:

S: 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.

There is a tendency for teachers, when confronted with an apparently incorrect statement, to inform the student of the error and perhaps state the correct point of view. For example, one can imagine at some point in the above discussion telling the student, "No, 50% would mean that the forecaster would expect rain on 5 out of 10 such days, just as 70% means that rain would be expected on 7 out of 10 such days." However, if a student's statements do not reflect isolated beliefs but are grounded in a more general conceptual framework, then one must question the effectiveness of these types of local interventions; they do not get at the heart of the problem. An analogous situation would involve correcting the assertion that a ship would fall off the earth if it ventured too far from shore by negating the assertion or citing evidence to the contrary rather than focusing on the apparent underlying belief in a flat rather than spherical planet. By understanding the outcome approach and the nature of various judgment heuristics, the teacher can "see beneath" superficial features of various student claims.

2.2 Statements That Sound Incorrect

The following is an example of a claim that seems patently incorrect. The conversation with a high school student who was participating in a summer mathematics program at Mount Holyoke College occurred spontaneously about half way through a two-week workshop on probability:

S: I don't believe in probability.
I: Why?
S: Because even if there is a 20% chance, it could happen.
I: Yes.
S: Even 1%, it could happen. I don't believe in probability.

On first reading, this student's claim seems baffling. If the assumption is made, however, that this student uses the term *probability* in a way consistent with the outcome approach, then it ought to be given the approximate meaning, "predicting single trials." Indeed, if that phrase is substituted for the word *probability*, the claim becomes not only comprehensible but normative. My interpretation is that this student is beginning to question the validity of her prior outcome-oriented view of probability. She is saying that she no longer believes that small probability values can be regarded as zero, and therefore one cannot predict with certainty the outcome of a chance event. She isn't yet aware, however, that the word *probability* can refer to something other than single-trial predictions.

It should be pointed out that the outcome approach is reasonably consistent with one of the common meanings of the related word *probable*. A probable event is, in everyday usage, one that is likely to occur. For example, George Shultz was reported in *The New York Times* (October 1, 1987) as saying, "I think there's not just a possibility, I think there's a probability," in response to a question about the likelihood of an eventual arms embargo against Iran by the Soviet Union. When asked to define probability, students often provide definitions of the type, "Probability is the chance of one thing occurring instead of another," or "Probability is what is most likely to be." This usage of the word *probability* to imply that some event is more likely to occur than not, fits well with students' outcome-oriented belief that mathematical probability involves *deciding whether* an event will occur, rather than *quantifying how often* that event will occur.

2.3 Two Views in Conflict

Although students can often fit what they read or hear in the classroom into their preconceived notions of probability, conflicts frequently arise between their view and the view being advocated by the teacher. These conflicts are welcome by the constructivist teacher. In the attempt to resolve the conflict, students often alter their understanding, bringing them into closer agreement with the normative perspective. An awareness of the student's beliefs permits the teacher to understand the nature of the conflict and possibly to help the student to navigate successfully. In the first excerpt discussed above, the student was experiencing some conflict over the appropriate unit of analysis in deciding the accuracy of weather forecasts. She resolved the conflict not by modifying her notion of probability but by reaffirming her belief that individual predictions are the appropriate unit. The excerpt below involves a student who experienced a similar conflict in responding to the probe about the accuracy of weather forecasts. He could not, during this conversation, come to a resolution, but moved back and forth between viewpoints in what he called a "logic swirl."

S: 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: What are you thinking?

S: Well, 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.

S: ... 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?

S: Well, that's weird cause it almost seems that ... 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 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.

The interaction described below took place during a lab on probability that was being run in conjunction with a remedial-level mathematics course at the University of Massachusetts. This particular episode is interesting both because of the extreme attempt made by the student to maintain his single-trial interpretation and also because by the end of the semester he had formulated a normative interpretation. He referred to this episode as being particularly important in the development of his understanding of probability.

The lab activity involved students guessing at the probabilities associated with a thumb tack landing with the point up (U) or the point down (D). For most thumb tacks, the probability of U is a bit greater than D. After making a guess as to the probabilities, each student conducted and recorded the results of 100 trials with the tack and then used these results to evaluate and, if necessary, revise the probability estimates. This particular student had initially guessed that the probability of U was .70. He then conducted 100 trials, and in spite of the fact that he had obtained 55 Us and 45 Ds, he concluded that D was more likely. I asked:

I: Suppose you had to bet on whether you'd get more Ds or Us in another 100 trials. What would you bet on?
S: I'd bet on D.
I: Why?
S: Because I got more Ds when I dropped it.
I: According to your results you got 55 Us and 45 Ds.
S: [Long pause while he verifies this result.] Well, ya. But I think D is more likely.
I: Can you explain why you think that, even though you got more Us?
S: Because it seemed like D came up more often when I dropped it once.
I: Which would be more likely if you dropped it 100 times?
S: OK. Well I guess U is more likely if I do it 100 times, but if I did it just once, D would be more likely.

This student eventually saw the inconsistency in his reasoning, but his last effort to salvage his belief that D was more probable by adopting opposite predictions for a single trial vs. 100 trials is a dramatic demonstration of the preference for thinking about probability in terms of single trials.

After resolving this conflict, the student became curious about why, having begun thinking U was more likely, he ended up thinking D was more likely even though he got more Us. He finally noticed that the first few trials were predominantly Ds and remembered that this had surprised him. He reasoned that thereafter he had "paid more attention" to trials resulting in D than those resulting in U. Forgetting for the moment whether his self-analysis was accurate, we see that his explanation is quite sophisticated and is consistent with an effect that Tversky and Kahneman (1973) have described as resulting from the "availability heuristic." According to this heuristic, the probability of an event is estimated by assessing the relative ease of bringing that event to mind. Thus people may determine the probability of winning a lottery by trying to recall people they know, or know of, who have won. Presumably, this is one reason that state lotteries and sweepstakes like to advertise using the names and homey photos of past winners – to provide us with a set of instances of winners whom we sort of know or perhaps can imagine knowing. One of the factors that can influence the ease of recalling event-instances is the vividness or saliency of an event. Witnessing a horrible traffic accident can, for a while, elevate a person's subjective sense of the probability of being in an accident. This student's account of why he came to think that the thumb tack was more likely to land with the point down appears consistent with the above account. Seeing what he didn't expect led him to "pay more attention" to subsequent occurrences of the same type such that his memory of the frequency of that event was distorted.

III. A FITTING APPROACH TO INSTRUCTION

I have suggested that students come into their first course on probability with a reasonably coherent and deeply engrained point of view about probability. One component of students' prior knowledge about probability involves judgment heuristics that operate much like visual perceptions. Thus, HTTHT just *looks* to most people to be a more likely result of flipping a fair coin than HHHHH. In addition to these perceptions, students also have a different conceptual understanding of the goal in situations involving uncertainty, i.e., that the goal is to predict individual outcomes. Together, these cognitions make learning the normative view difficult.

Although these informal models and methods pose a serious challenge to those trying to teach and learn probability, there is emerging evidence that instruction in misconception-rich domains can be effective. In physics, where students also have strong prior conceptions, it has been suggested that it may be effective to encourage students to recognize and resolve conflicts between normative concepts and erroneous intuitions (McDermott, 1984). Several researchers, including Minstrell (1982), Clement (1987), and Hake (1987), have demonstrated that physics instruction specifically designed to address various misconceptions can be effective. Their approach includes laboratory exercises designed to demonstrate counter-intuitive results and promote student discussion, problems that require qualitative rather than quantitative solutions, as well as presentations that explicitly contrast normative with non-normative physics concepts.

Inspection of the various classroom techniques being used to overcome physics misconceptions reveals three general criteria against which students are encouraged to evaluate their current beliefs: the fit between their beliefs and 1) the beliefs of others, 2) their other, related beliefs, and 3) their own observations. I conclude this chapter with a brief discussion of these criteria with the conviction that if instruction in probability is to be effective, the teacher will need not only to be attuned to student intuitions but to structure activities that encourage students to evaluate those intuitions in accord with these three criteria.

1. *Do my beliefs agree or fit with the beliefs of others?* A typical starting point for the presentation of some physics concept or principle is to invite students to discuss among themselves what they believe about a particular situation. For

example, Minstrell (1982) asks students to determine the forces that are acting on a book lying on a table. After allowing some time for students individually to think about the question and to draw a diagram showing all of the forces acting on the book, he has students share their answers with the class. This sharing leads to a dialogue, with students questioning and challenging one another. It should be pointed out that discussions that allow for the expression and defense of differing points of view are not common in the everyday experiences of the students. They are certainly not the norm in the classroom nor do they occur in typical social situations. As Stubbs (1983) observes,

> ... there is a general rule in our society that demands that interaction proceed at a smooth flow: silences are often considered embarrassing and disagreements must normally be mitigated. So speakers immediately counteract departures from the smooth ongoing of normal face-to-face interaction by making (if necessary, violent) attempts to restore the ritual equilibrium (p. 241).

Thus, even though instructors may say little during this discussion, they play a critical role in keeping the conversation going, being supportive of students' expression of their views, and at times helping to focus the discussion, but above all, creating an atmosphere in which it is acceptable to articulate an opposing view and to challenge what one does not understand or believe.

The class discussion accomplishes a number of goals. First, it allows students to make explicit to themselves what they believe about a situation before they know what the "expert" view is. Second, classroom discussion provides a motivation for further exploration of the question. The act of articulating one's beliefs in a public forum involves a personal investment in the question. Students who express their view are no longer indifferent to the final outcome of the discussion – in general they want to have been "right" rather than "wrong." Third, discussion among students provides the opportunity for the instructor to gain further insights into how students are thinking about a particular topic. This information can be used to plan future interventions and to monitor conceptual development. finally, the discussion communicates to the student the value that the instructor places on students' understanding and expression of ideas.

Once students have made their own beliefs more explicit, the second criterion can be applied:

2. *Are my beliefs internally consistent?* In the process of trying to convince one another, students explore in greater depth the implications of and interconnections among their own beliefs. At times this exploration can lead students to discover inconsistencies in their beliefs or gaps in their reasoning. For example, Minstrell (1982) asks a series of questions related to the book-on-the-table problem mentioned above. Roughly the same number of students who initially do not believe that a table exerts an upward force on a book resting on the table, do not believe that an upward force is exerted on a book resting on an out-

stretched hand. However, as more books are added to the outstretched hand, more students subsequently express a belief in an upward force, and the majority of students believe that there is an upward force on a book resting on a spring. As the students progress through these questions, they are repeatedly asked about the book on the table, and an increasing number indicate a belief that the table pushes up on the book. The students are apparently motivated to see the situations as analogous and therefore to be consistent.

In addition to asking students to extend their analysis to a variety of related instances, various probes can be effective in encouraging students to analyze the consistency and completeness of their beliefs. These include probes of the form,

Why do you think that ...?
How would you explain a situation where ...?
How does that relate to what you said earlier?
What would you say to a person who gave the following argument ...?

One reason that misconceptions are so difficult to alter is that they tend to comprise a coherent, self-consistent framework. To the extent that a person's beliefs are self-consistent, they are impervious to this type of challenge. This is one reason why the third criterion is especially important.

3. *Do my beliefs fit with empirical observations?* Beliefs about physical, social and probabilistic phenomena are always to some extent based on or confirmed by observations of events. This includes beliefs that are regarded by the "educated" as irrational: People who believe in astrology will argue that predictions based on their astrological chart have come true, and paranoid schizophrenics will cite numerous observations to support their claim that others are conspiring against them. The fact that what one observes is never independent of what one already knows or believes, means that data cannot serve as the final arbiter of so-called "empirical questions." How is it then that the fit between empirical observations and beliefs can be considered a criterion in evaluating those beliefs? This is a complex question which I cannot discuss adequately here. But there are two things that people generally do not do that further weakens the informative power of empirical observations: 1) they do not keep accurate records and 2) they do not look for data that would be inconsistent with and thus controvert, a belief they hold. To be effective, classroom demonstrations meant to challenge students' beliefs should emphasize these two features.

Although demonstrations have been a traditional part of physics instruction, they seldom involve students first speculating and making predictions of what they think will happen based on their own understanding of the situation. Minstrell (1982; 1984) has used demonstrations along with this type of discussion and prediction very effectively. With the arrival of the computer in the classroom, similar demonstrations of probabilistic and statistical phenomena are now practicable. Large random samples can be quickly and repeatedly drawn and summary results computed and displayed. By having students first make predictions of what they expect to observe, such simulations can be used to challenge non-normative beliefs about random events.

The three criteria for evaluating current beliefs are appropriate for use in the constructivist classroom not only because they are effective in inducing conceptual change but also because they are the same criteria that define scientific activity. Scientific theories that are worth further consideration must first be made explicit, they must not contain internal contradictions, and they must fit with observations. Students with strong prior conceptions are therefore treated in this instructional approach as scientists, in as much as they are viewed as holding theories which they should not be expected to abandon until these theories have proven unsuccessful (and not just unsuccessful vis-à-vis course grades). Furthermore, the testing of the adequacy of student beliefs should not be an isolated, individual enterprise but, as in scientific activity, involve a dialogue that takes place among a community which negotiates the relative merit of various perspectives.

REFERENCES

Cherry, C. (1968). *On human communication*. Cambridge, MA: M.I.T. Press.

Clement, J. (1987). Overcoming students' misconceptions in physics: The role of anchoring intuitions and analogical validity. In J. D. Novak (Ed.), *Proceedings of the Second International Seminar, Misconceptions and Educational Strategies in Science and Mathematics* (pp. 84–97). Ithaca, NY: Cornell University.

Fisher, K. M., & Lipson J. I. (1985). Information processing interpretation of errors in college science learning. *Instructional Science, 14*, 49–74.

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

Hake, R. R. (1987). Promoting student crossover to the Newtonian world. *American Journal of Physics, 55*, 878–884.

Kahneman, D., & Tversky, A. (1972). On prediction and judgment. *Oregon Institute Bulletin, 12*(4).

— (1973). On the psychology of prediction. *Psychological Review, 80*, 237–251.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction, 6*, 59–98.

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

McDermott, L. C. (1984). Research on conceptual understanding in mechanics. *Physics Today, 37*, 24–32.

Minstrell, J. (1982). Explaining the "at rest" condition of an object. *The Physics Teacher, 20*, 10–14.

— (1984). Teaching for the understanding of ideas: Forces on moving objects. In C. W. Anderson (Ed.), *Observing science classrooms: Perspectives from research and practice* (A.E.T.S. Yearbook XI). Columbus, OH: ERIC Center for Science, Mathematics and Environmental Education.

Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Bulletin, 68*, 29–46.

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance*. New York: Norton. (Original work published 1951)

Stubbs, M. (1983). *Discourse analysis: The sociolinguistic analysis of natural language*.

Chicago: University of Chicago Press.
Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Review, 76*, 105–110.
— (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.