17. Dimensions in the Assessment of Students' Understanding and Application of Chance

Kathleen E. Metz

Purpose

   Randomness and chance variation are key ideas that can function as goals in young students' understanding and application of chance. In this chapter I examine how these key ideas involve construction of new concepts, as well as beliefs about the place of chance in the world. These ideas are considered from the perspective of the mathematics or statistics classroom culture; i.e., how the classroom culture reflects and fosters beliefs about the place of uncertainty and chance in the world.

IDENTIFICATION OF KEY IDEAS

   The National Council of Teachers of Mathematics' *Assessment Standards for School Mathematics* (1995) asserts that current classroom assessment practices frequently tap the "uninteresting and superficial." According to this document, we need to fundamentally reframe our assessment tools, so that they reveal students' knowledge of core ideas and how students apply these core ideas to new and familiar problems. This criticism of contemporary assessment practices parallels NCTM's (1989) concerns, delineated in the *Curriculum and Evaluation Standards for School Mathematics*, that contemporary instructional practices pay too much attention to algorithms and too little attention to the concepts and big ideas underlying the algorithms.
   Instruction and assessment in statistics and probability have frequently constituted an extreme example of a focus on procedures to the neglect of underlying concepts and big ideas. In this vein, Stodolsky argues (1985) that even university students are frequently taught statistics with the assumption that experts will tell these students what to do if and when they subsequently need statistics for their own purposes; thus conceptual understanding appears neither expected nor necessary. One important step in the strengthening of instruction and assessment in the domain of statistics and probability involves identification of concepts and big ideas as a key frame in the analysis of the conceptual underpinnings of the curricula and the adequacy of students' evolving knowledge in the domain.
   The national curricular reform documents of Australia, Great Britain and the United States each recommend that students in the first level of schooling reflect on ideas of chance and probability, within the context of interpreting data they have collected themselves (Romberg,

Allison, Clarke, Clarke & Spence, 1991). Analysis of the psychological and mathematical literature concerning chance reveals the deep complexity and multifaceted aspects of these constructs and their application. For the purposes of more specifically identifying the big ideas in this sphere that are within reach of young students, we find extremely useful the thinking of statistician and educator, David Moore (1990). He points us to two constructs, randomness and chance variation, that he argues can be fostered through appropriate childhood experiences.

Randomness

Both the cognitive and mathematical literature manifest considerable disagreement on the meaning and entailment of randomness. For the purposes of assessing (young) students' curriculum and emergent learning, we use a basic definition offered by Moore. Moore explains:

> Phenomena having uncertain individual outcomes but a regular pattern of outcomes in many repetitions are called *random*. "Random" is not a synonym for "haphazard," but a description of a kind of order different from the deterministic.

Note that the construct of randomness and its attribution to a particular phenomena assume two aspects: the uncertainty and unpredictability of the single event, in conjunction with the patterns that emerge across a large number of repetitions of the event. Thus, although certainty of prediction is unattainable for individual random phenomena, at least prediction becomes possible over the long haul.

Chance variation

A major issue that emerges in any context of data interpretation is the question of whether the variations in the data have some kind of causal base or are simply due to chance. This issue can emerge as young students consider what the data they have collected do and do not mean, as well as in the more sophisticated interpretations of older students. Adequate interpretation of data demands an understanding of the idea that variations can be due to either chance and/or some form of underlying causality. Moore contends that it is crucial for students to grasp the idea that "chance variation, rather than deterministic causation, explains many aspects of the world" (Moore, 1990, p. 99). This key idea is fundamental to students' negotiation of data-based curricula, as well as their adequate interpretation and prediction of patterns outside of school — from accidents to sequences of shots on the basketball court.

THE CHALLENGE OF ASSESSING STUDENTS' UNDERSTANDING AND APPLICATION OF CHANCE

Assessment of students' understandings poses considerable challenges for teachers. Most simply, many kinds of tasks used in the instruction and assessment of statistics and probability may reflect little about students' underlying conceptual knowledge, but rather their ability to apply the right algorithm. For instance, consider an exemplar of a statistical activity for grades 3 - 4 given in NCTM's *Curriculum and Evaluation Standards*:

"[Given a spinner divided evenly into four sectors, with three sectors colored red and one sector colored blue] "Is red or blue more likely? How likely is yellow? How likely is getting either red or blue? If we spin twelve times, how many blues might we expect to get?" (NCTM, 1989, p.56)

In this kind of problem, right answers could easily hide students' failure to grasp how chance enters into probabilistic situations. For example, in relation to the last question posed, the answer of "9 reds" may well reflect analysis of the ratio of colors in the spinner. However, this response does not reveal whether the student has grasped the idea that although 9 is indeed the best prediction, this particular outcome is uncertain. Without a keen attention to the reasoning underlying students' predictions— as revealed in class discussions or written work— the teacher would have little information about students' understanding of chance.

Above and beyond the significant conceptual challenges of this domain, appropriate application of chance also involves beliefs. Whether or not students will think to interpret a situation in terms of chance will be influenced by their individual beliefs about the place of chance and uncertainty in the world (Nisbett, Krantz, Jepson, & Kunda, 1983). Furthermore, the individual's beliefs about appropriate data interpretation and how chance may enter in are also influenced by the classroom culture and the beliefs and values it reflects and encourages.

In an effort to address these complexities, this chapter analyzes assessment of the understanding and application of chance along three dimensions: *(a)* cognitive: the student's conceptual constructions; *(b)* epistemological: the student's beliefs about the place of chance and uncertainty in the world; and *(c)* cultural: the culture of the mathematics classroom and learning environment vis à vis beliefs about how chance and uncertainty fit in and related dispositions of interpretation and inference. Each dimension is examined through consideration of its relevance, analysis and illustration of relatively naive versus more adequate stances, and its implications for assessment. Illustrative examples are based on the authors' research and work with children in the early grades, but the points raised are general and for the most part apply to older students as well.

THE COGNITIVE DIMENSION: CONCEPTUAL CONSTRUCTION

A more conceptually-oriented approach to students' study of chance and their interpretation of variability in data demands assessment of students' conceptualizations of the core idea of randomness and the alternative interpretations that students may evoke instead. The manner of teaching now encouraged by the mathematics reform movement — including fostering of mathematical discourse among students and critical analysis of the meaning of data they collect — manifests a range of rich indicators of students' thinking that can reflect the distinctions in students' understanding of chance described below.

The challenge of understanding randomness

As noted above, randomness involves both the uncertainty and unpredictability of a single event, with an understanding of the patterns that can emerge over a long sequence of repetitions of the event. The challenge of integrating uncertainty and pattern aspects into the randomness construct constitutes a primary issue in cognitive development and a primary source of errors or "bugs" among older students and adults.

The cognitive literature has documented young students' attribution of uncertainty and unpredictability to what are actually random phenomena, without a corresponding sense of the patterns involved. Conversely, it has also documented young students' attribution of deterministic patterns to random phenomena, without any sense of the corresponding element of unpredictability and uncertainty. While a *patterns-without-uncertainty* interpretation exaggerates the information given, an interpretation of *uncertainty-without-patterns* underestimates the information given. One fundamental issue to consider in the assessment of students' grasp of randomness is the status of the construction; i.e., evidence of the randomness construct in integrated form versus evidence of an interpretation limited to patterns or uncertainty.

Extending the construct of randomness across probability-types

Probability is not a singular construct, as evident explicitly in the perspective of the expert and implicitly in the young student's various ways of deriving and justifying probabilities. The focus below is on the two types of probability that most frequently arise in the current approaches to teaching students chance and probability: *classicist probabilities* and *frequentist probabilities*. We have no reason to believe that the different ways of deriving and conceptualizing probability pose equivalent challenges to students. Furthermore, we have no basis to assume that students' understanding of one form of probability entails any understanding of other probability-types.

Classicist probabilities

The use of chance-generating devices, such as die, coin tossing, or spinners constitutes a common instructional technique for introducing the idea of randomness. Classicist probabilities denote the derivation of probabilities from an analysis of the symmetries in chance-generating devices. For example, as a coin has two sides, it has a probability of .5 of landing on tails and .5 of landing on heads. Indeed, such games, analyzed in terms of their fairness, appear early in many probability curriculum materials for children and adolescents. This approach enables relatively efficient data collection and relatively straightforward derivation of expected probabilities, in the form of classicist probabilities.

Frequentist probabilities

This type of probability denotes the frequency of a given outcome over an infinite number of repetitions of the event. It can be approximated by conducting experiments that involve repeating events many times, and noting the relative frequencies that emerge over many trials. Interpretation of these data and derivation of probabilities would not be amenable to a classicist perspective, but will instead draw on children's intuitions of frequentist probabilities. (As noted above, US, English and Australian mathematics education documents recommend that children conduct simple investigations about themselves and the world around them.)

As frequentist probabilities presume some understanding of the notion of infinity, this probability-form may be particularly challenging to young children (Hawkins & Kapadia, 1984). There is evidence that aspects of the construct of infinity are difficult for school-age children (Piaget, 1987; Nunez, 1993). Nevertheless, children can grasp the precursor construct of probabilities in terms of the distributions that emerge across many repetitions.

ANALYSIS OF STUDENTS' COGNITIVE CONSTRUCTIONS.

The following sections focus on the analysis of student's cognitive constructions, from the perspective of the status of the randomness construct and probability-type. First, examples of patterns without uncertainty are considered, followed by uncertainty without patterns, and finally uncertainty with patterns combined in the construct of randomness. Each of these cognitive constructions is illustrated in the form of classicist and frequentist probabilities.

Patterns without uncertainty

Students can reason in terms of patterns, but fail to conceptualize the chance involved in these patterns. Hence they exaggerate the information given. The interpretation of patterns without uncertainty in classicist form is illustrated through a child's analysis of the effect of different spinners on game outcomes. In this example, the child derives absolutist predictions of outcomes from the symmetry of the spinner-device:

The teacher shows Mary, a third grader, two spinners. One of the spinners is evenly divided into two sectors, one red and one yellow. The other spinner is 3/4 red and 1/4 yellow. The teacher explains that a player gets to advance up the game-board when the spinner lands on their color during their turn. She then asks Mary, "Do you think it will make any difference which spinner we play with?" Mary argues that it will make a difference, because when you play with the spinner with the even division of red and yellow, "Sometimes you turn it, it'll stop on the red and sometimes you turn it it'll stop on the yellow."

She denies this uncertainty applies to the other spinner with the asymmetrical color division. When the teacher asks "When you turn this [3/4 red spinner; 1/4 yellow] does it sometimes stop on the yellow and sometimes stop of the red?" She shakes her head. "No." When the teacher asks her to predict where the player with the yellow color will be at the end of the game, Mary predicts the board starting position.

This excerpt reflects patterns without uncertainty, in that the student infers that spatial predominance determines each outcome (as indicated by an analysis of the symmetries in the outcome-generating device). Less primitive forms of patterns without uncertainty involve the inference that the outcome associated with spatial predominance will always win overall, but the outcome associated with a subordinate spatial value will occur periodically and any particular spin is uncertain.

In terms of the challenge of assessment, note that prediction of who would be most likely to win the game would not have differentiated Mary's buggy construct from the adequate randomness construct: both interpretations would have predicted the predominant color as the winner. The inadequacy of Mary's conceptualization of chance is manifested in her predictions about the placement of the losing player, and more generally, the certainty which she attributes to all of her projections.

Patterns without uncertainty as a precursor to frequentist probabilities is illustrated through excerpts of a student's approach to a sampling problem. In this situation, the student views the distribution of elements in the small sample drawn as an accurate reflection of the urn contents. The missing aspect of uncertainty takes the form of the student's failure to recognize how chance enters into the sample viewed. More generally, this manner of thinking reflects what Tversky and Kahneman (1971) have facetiously called the Law of Small Numbers:

> The teacher shows an urn, filled with a hidden collection of marbles to Janice, another third grader. She asks Janice if she can figure out what's in the urn without dumping them all out and without looking. Since Janice has no idea how to proceed, the teacher suggests she iteratively take one out, look at it, and replace it. The hidden collection actually has an even number of red and blue marbles. She supplies Janice with a collection of chips the same colors, to help her keep track of her marble draws.

> Janice makes 12 draws with replacement, each time adding another red or blue colored chip to her piles to keep track of her draws. She concludes, "I think it's mostly red, because there's seven of the red and 5 of the blue." When she checks her conclusion by peeking into the jar, she looks astonished at her discovery that there are not more reds. Janice explains that she had thought there were more reds, "Because there was more red [chips] than blues [chips] on the table". Upon probing, Janice denies there is any way she could have figured out there was the same number or any way she could have been more sure of her prediction.

In this case, the student's patterns without uncertainty interpretation is reflected by her confident interpretation of the meaning of small differences in the subsets of the sample, her bewilderment when she discovers her inference is wrong, and her failure to consider the potential of extending the action that the teacher viewed as the sampling procedure.

Conceptualizations of patterns without uncertainty are reflected in a number of buggy interpretations identified even at the adult level. The Gambler's Fallacy or negative recency affect denotes the expectation of local corrections to random fluctuations in a sequence; e.g. the assumption that the toss of two tails in a row will be followed by the toss of a head. Vallone and Tversky (1985) have identified a related bug, where the individual assumes random fluctuations in the data must be causal and proceeds to develop causal explanations.

Uncertainty and unpredictability without patterns

Whereas in patterns without uncertainty, students assume they can confidently predict what will happen, thereby exaggerating the information given, in *uncertainty without patterns* students conceptualize the situation as simply unpredictable, thus underestimating the information given. In the situation with the spinners, uncertainty without patterns is manifested by the belief that as long as both color outcomes are included in the spinner, the situation is impossible to predict:

> The teacher asks Karen, a kindergartner, if she thinks it'll make any difference which spinner they play with. Karen argues, "No," 'cause they both have yellow and both have red." The teacher suggests they play with 3/4 red, 1/4 yellow spinner and invites Karen to choose her color. She selects yellow for herself, the color of the bad odds, and has no predictions about who might win.

> After the child has lost the game, the teacher says, "I want you to look real carefully and see if you think it makes any difference which one you play with." Karen maintains her position: "Nah, it wouldn't because they [the spinners] both have yellow and they both have red." "Do you think it makes any difference if there is more red on one than the other?" the teacher hints. "No", Karen argues, "Because they have yellow and they both have red."

In this example, Karen's buggy conceptualization of uncertainty without patterns is reflected in her choice of the spinner with the bad odds for her color, her absence of predictions for the game outcome and, more generally, her conviction that it is irrelevant which spinner they play with, as long as the spinner has both players' colors.

The failure to anticipate the differential effect of asymmetries in the chance-generating device is relatively primitive. Most kindergartners do anticipate differential effect, although typically without a corresponding appreciation of the chance involved. However, uncertainty without patterns is manifested by much older students and even well-educated adults in connection with frequentist probabilities.

Both children and adults reflect this manner of thinking in the urn sampling problem. In contrast to the spinner situation, many sampling problem-types involve a sampling space initially in the form of an unknown. Children and adults alike can focus on the collection in the urn as an unknown and fail to appreciate how a succession of draws can inform them about its contents:

> Third grader, Nathan, continues to be skeptical about the feasibility of the urn sampling task, even after the teacher suggests the procedure of drawing and replacing a marble at a time: "I don't think that'll work because I might like have taken the same marble out a couple times." "Do you think if you did it a *whole* bunch of times you could figure it out?", the teacher suggests. "No," Nathan insists. "It's impossible."

In his acute awareness of the uncertainty in the situation, Nathan has failed to realize that patterns can emerge across extended repetitions of the random event and that these patterns do give us some information, although imperfect, about the population from which they were drawn.

The construct of randomness: learning and assessment

Application of the construct of randomness demands an integration of the uncertainty and unpredictability of a single event, with an understanding of the patterns that can emerge across repetitions of events. The cases cited here were from students who had manifested considerable learning on task. The first example is Karen, who had previously argued that, as long as both players' colors were on the spinner, the game was unpredictable. Consider her reasoning four games later:

> In conjunction with the two-colored 3:1 red/yellow spinner, Karen chooses the red chip for herself, relegating the yellow chip to the teacher. Karen explains she wants to play with the red chip, "Cause the red has more than the yellow. Because if you play with the red, you have a very good chance of winning."

Karen's grasp of randomness is reflected in her choice of the chip with the better odds, in conjunction with her explanation of its affect: her probable but uncertain win. Her derivation of

differential probabilities is based on an analysis of spatial relations in the chance-generating device.

Similarly, recall that Nathan had previously contended that prediction of the urn contents was impossible simply on the basis of single draws. With scaffolding, Nathan begins to integrate notions of uncertainty and unpredictability with the idea that patterns that can emerge across large numbers of draws. He remains convinced that, no matter how long one continues sampling, you can never know for sure:

> Although skeptical of its utility, Nathan agrees to try the procedure of one-by-one sampling with replacement that the teacher suggests. After drawing 7 marbles, including 6 reds (which Nathan considers orange) and one yellow, he comments, "I think they're more of the oranges now because I keep on taking out oranges." "Are you quite sure?" probes the teacher. "Not really," he responds.

> When the teacher asks Nathan if it would make any difference if he continued the draw and replacement procedure, he immediately picks up on the idea, exclaiming, "I think it would give me a better idea!". After 8 more draws, consisting of 7 oranges and one yellow, Nathan concludes, "I think it's more orange." "Are you quite sure?" questions the teacher. "Uhhuh," he affirms. "Any way you could be very sure or is that impossible?" she probes. "That's impossible!" exclaims Nathan.

Note that Nathan's prediction ("It's more orange") could stem from either a deterministic interpretation of patterns or an integration of chance and patterns. It is his qualifications concerning his confidence in this prediction, together with his assertions that certainty is unattainable, that indicate the integration of chance and patterns in the randomness construct.

When teachers focus classroom discourse on such issues as the meaning of patterns and variability in the data, and the bounds of predictability and control, they can simultaneously support the learning and assessment process. Teachers can use these discussions to analyze the students' grasp of randomness and chance variation. These constructs are fundamental in their learning to successfully cope with uncertainty, in experiences beyond the school walls as well as their formal study of statistics.

THE EPISTEMOLOGICAL DIMENSION:
BELIEFS ABOUT CHANCE AND RANDOMNESS

The differentiation of deterministic from nondeterministic events challenges us all; children and adults, novices and experts. Indeed, the demarcation of the causal from the stochastic stands at the forefront of many fields, including genetics, physics and ecology. Analysis of intellectual history (Gigerenzer, Swijtink, Daston, Beatty, & Kruger, 1989; Hacking, 1990) and contemporary bugs in adults' thinking (Kahneman, 1991; Kahneman, Slovic & Tversky, 1982) reveal a general tendency to err toward the side of attributing too much to deterministic causality, with a corresponding failure to recognize the extent to which chance operates in what one experiences of the world.

If students are to appropriately use the interpretative frame of *chance variation*, they need to both grasp the idea of randomness *and* have a view of the world with a place for chance events. In other words, a student may have grasped the idea of randomness, but unless she believes that some phenomena are indeed due to chance she will not apply the construct. This epistemological

dimension of chance is relevant to teachers, in that variations in epistemological stance of this kind can affect students' ability to grasp these big ideas, as well as their propensity to apply chance interpretative schemas. This chapter uses the term *epistemological set* to denote an individual's predilections to interpret the world in relatively deterministic or stochastic terms, based on their beliefs about the place of chance in the world.

Epistemological set does not imply that individuals do not have both chance and deterministic interpretations within their cognitive repertoire, but that they have a tendency to interpret phenomena toward one or the other end of the stochastic / deterministic continuum. As Fischbein argued, "What we are really and specifically concerned with here is the *orientation of knowledge*, a factor of cognitive 'set' or a kind of 'mentality' on the intellectual level." (Fischbein, Pampu & Minzat, 1975, p. 170).

This chapter examines students' epistemological set, as a second dimension in teachers' assessment of students' understanding and application of chance. The sections below consider examples from students with epistemological sets from both ends of the continuum and the nature of the indicators that differentiate them.

Students with relatively deterministic epistemological sets

Data interpretation of two different students are used below as illustrations of relatively deterministic epistemological sets. The first case is Christy, a six year-old kindergartner. Her data interpretations are reported from two tasks. One task, adapted from Piaget and Inhelder (1975) involves predictions and interpretations of the mixing of marbles, initially arranged by color, in a box that can tilt up and back. The other task is the urn sampling problem described above.

> The teacher shows Christy a shallow box affixed on an axis. Along the lowered edge of the box are 12 marbles, including 6 black ones on the right and 6 silver ones on the left. The teacher asks whether or not the marbles will return to their original positions after the box is tilted back and forth. After initial predictions about movements of the marbles upon rotation of the box up and back, Christy takes the box from the teacher, telling her she wants to figure out what's happening. Christy exclaims, "They're trying to get back to their places! These [black marbles] want to get over here and these [the silver marbles] want to be right here! " The teacher probes, "Why do they want to be in those places?" Christy explains, "Because that's the way God made them!"

Christy's assumption that the marbles are trying to return to their places reflects a view of the world as ordered. It doesn't occur to her that the marble movements may be affected by chance factors.

In the urn sampling problem, Christy's interpretations again manifest a relatively deterministic epistemological set. Here her deterministic epistemological set is manifested first through the topology of the bottom of the urn that she argues determines which marble is drawn next; and subsequently through the potential control one has over marble color drawn.

> After just 4 draws, Christy is confident that there are more greens than reds. "Are you quite sure?" probes the teacher. "Yah!" she immediately responds. Christy then inspects the urn contents to discover the same number of reds and greens. The teacher asks her, "Why did you think there were mostly green when actually there were the same number?" In response, Christy appears to assume a

kind of pathway in which the marbles are arranged by color and thus proceed one color before the other: "Because it was like this. Every time I would drop it in front, so I would keep on getting green until I got a red. Maybe when I picked up, I always picked up green, green, green green, until the reds finally came ahead."

Finally, Christy suggests that she could have gotten it right by strategic marbledraws: "They could maybe go different by different. First go red, green, red, green." The teacher questions, "But how would you be able to do that?" "Maybe if I pulled in different directions," reasons Christy.

In her pathway model of marble selection, the selection process would be largely determined. Of course, this model totally ignores the random nature of the arrangement in the urn. Her idea that she should have been able to draw the marbles in an alternating order, such that the equality of the different colors in the urn was transparent, also fails to appreciate the essential random nature of the draw.

The second example of a relatively deterministic epistemological set is drawn from one child's thinking within a class-based investigation, suggested in the Used Numbers book, *Statistics: Middles, Means and Inbetweens* (Friel, Mokros, & Russell, 1992). A fifth grade class has been collecting, representing, and interpreting data concerning the number of raisins in a box. The teacher has asked the students to formulate their best guess concerning the number of raisins in a box, given data about the contents of 28 other boxes they have already opened.

Building on a representation of the data invented by a small group of students, the teacher has organized the class data into a line plot. She shows the students the line plot and then challenges them to figure out what would be the best prediction for the number of raisins in an unopened box. Sue's strong prediction is 72. She argues, "There is no 72 [in the distribution to date]! And it's really close to the average [the mean has been calculated as 73]. You kind of wonder why there isn't 72. You have to wonder why there isn't a 72. I think the next box will have 72."

It doesn't occur to Sue that the absence of any 72 data point could be due to chance. In assuming that the dip in the curve will be filled by the next data point, Sue's reasoning reflects a spurious determinism. She fails to grasp the fact the each raisin box sample is an independent event, not affected by the particular boxes that happen to have been counted previously. Her thinking is closely related to the common adult bug called negative recency effect or Gambler's Fallacy, in which one anticipates local corrections to random fluctuations. Again, the student's answer alone would not have revealed her deterministic set. This orientation becomes evident in her explanation of why she infers 72 would be the best answer.

Students with epistemological sets emphasizing chance and uncertainty

Two illustrations of epistemological sets emphasizing chance and uncertainty are considered below. The first case is Nathan, whose acute sense of the unknown and the uncertain was noted above in connection with the urn sampling problem. Nathan's thinking in relation to the marble tilt box (see Christy's case in the last section) and the spinner reflect a similar emphasis of chance and uncertainty, *albeit* in different forms:

In the marble tilt box problem, Nathan assumes some kind of unpredictable change in   the marble arrangement. He explains, " The marbles will lose direction. They will loose control." In response to the teacher's query of whether or not the marbles will return to   their original positions, Nathan will not commit himself beyond "Could be." When asked   how long this might take, he sticks to "I don't know."

In the spinners task, Nathan is guarded about all his predictions. For example, when   analyzing the affect of playing with the 3:1 red/yellow spinner, Nathan notes "I *think* with   that   color   [red]   I'd   win, because it's [the spinner] more of my color." Even when he is   way ahead of the teacher, also burdened with bad odds, he predicts he'll win, but also notes she still could win "If you keep on hitting on the yellow."

In my experience of interacting with Nathan across six different activities, he was consistently prone to interpret data or predict outcomes in non-deterministic terms. Chance and uncertainty appear to enter into Nathan's view of the world, from both his assumptions of how the world works as well as intense awareness of his own ignorance.

The second case is John, another third grade boy who also appears to view the world in terms emphasizing chance and uncertainty. John's perspective is briefly described in relation to the marble tilt box problem, analysis of the spinners, and the urn sampling problem.

John is reluctant to commit himself on a prediction concerning what the marbles might   look like after a rotation up and back, simply commenting "It might not [come back to the   black marbles on one side and silver marbles on the other]. The teacher's query, "Do you think eventually the black ones will be on the other side from where they started out?",   leads to a similar noncommittal response: "It might happen some time."

In the context of the spinner games, John spontaneously tells the teacher, playing with   the   yellow chip in conjunction with the 3:1 red/yellow spinner, " You have a chance. . . .
if you get lucky."

John is concerned about the feasibility of prediction of the urn contents, "Because every time you put it [the marble drawn] back, you might get the same color as before, so you   can't be sure."   The teacher suggests, "If you kept on doing that for a long time, would   you be pretty sure [if they're mostly one color or the same number of each]?" John   remains tentative, telling the teacher, "You never know."

John seems attuned to the uncertainty of events and, conversely, reluctant to make firm predictions on the basis of anticipated patterns. His epistemological set, emphasizing chance and uncertainty, is manifested in his reluctance to make predictions, his acknowledgment of the possibility of multiple outcomes, and his strong sense of the unknown and unknowable.

Implications for Assessment

Epistemological set — i.e. the students' beliefs about where and to what extent chance enters into the world — are relevant to the classroom teacher who teaches statistics, in that it will strongly affect the students' propensity to consider a stochastic interpretation. Epistemological

set can be reflected in the assumptions students make about causality and chance and their respective spheres of application. Differences will be manifested in discrepant ways in which students interpret and explain the same data set: seeing determinism in variability versus seeing probabilistic patterns or uninterpretable uncertainty. These distinctions in epistemological set can be reflected in whole class or small group discussions about appropriate data interpretation, where (as in the example of Sue's class noted above) students not only give their answers but discuss why they consider their interpretation the most reasonable. More generally, students' epistemological set is revealed in their propensity to assume deterministic explanations versus their willingness to seriously consider the possibility that the outcomes may be due to chance.

THE CULTURAL DIMENSION: HOW THE CLASSROOM ENVIRONMENT TREATS CHANCE, UNCERTAINTY, AND BELIEFS

I had to rethink my interpretation of Christy's remark that God made the marbles move the way they had after her principal asked me, "You don't *really* believe that anything is due to chance, do you? We teach our children that God is behind all things." To adequately understand students' cognitive constructions and beliefs, we need to consider the culture in which students participate. The extent to which individuals assume that deterministic causality underlies variability, as opposed to the possibility of random variation, depends in part upon the orientation of their society at large and school culture. According to Nisbett, Krantz, Jepson and Kunda (1983), the culture influences the kinds of situations in which individuals are likely to consider a stochastic interpretation.

We can conceptualize students' learning statistics and probability as in part a process of enculturation. Indeed, Lauren Resnick has argued that becoming a mathematical problem-solver is in large part a process of enculturation, "as much acquiring the habits and dispositions of interpretation and sense-making as of acquiring any particular set of skills, strategies, or knowledge" (Resnick, 1988, p. 58). One subtle yet important parameter of assessment in this domain concerns the extent to which the culture of the math classroom, in the activities it structures and the interpretations it values, embodies a deterministic versus nondeterministic view of the world. We can conceptualize this dimension of assessment as a kind of epistemological set at the level of the social group within the social institution of the school. Below we consider assessment from the cultural perspective of messages about the place of chance and determinism, implicit in the values and habits of the classroom learning environment.

The task of assessing the culture of the mathematics classroom in terms of epistemological set involves a shift in emphasis from learner to learning environment. Indicators of classroom epistemological set include the choice of subject matter, the structuration of problems, appropriate means of data interpretation, sufficient evidence to assume causality, and the aesthetics of what constitutes a good solution or explanation.

Classrooms that support a deterministic view of the world are easy to find. A very prevalent method of teaching students mathematics has been and continues to be teacher telling, followed by text-based students' practice (Goodlad, 1984; Stodolsky, 1985). Within this approach, there are right answers and wrong answers, with little ambiguity about what algorithms should be applied. Indeed, the school culture has been identified as, in general, a strong influence in the direction of a deterministic epistemological set. Moore (1990) points to the choice of curricular

emphases in the initial years of children's formal schooling as a factor in the building of a deterministic world view, particularly in the realm of mathematics. Moore writes:

> Children who begin their education with spelling and multiplication expect the world to be deterministic, they learn quickly to expect one answer to be right and the others wrong, at least when the answers take numerical form. Variation is unexpected and uncomfortable (Moore, *ibid.*, p. 135).

In addition to the fact that the basic skills curriculum tends to be strictly deterministic, where ambiguity does come in (e.g., the best representation to use in solving a word problem), the authority in the guise of teacher or text provides the single correct form.

More specifically, Fischbein (1975) argues that schooling tends to have detrimental effect on children's understanding of chance. Indeed, this researcher points to evidence in his work and in unexplicated findings in the work of others of a regression in children's grasp of randomness after the point of entry into formal schooling, a regression which he attributes to the influence of the school culture. Fischbein asserts:

> It is generally true that contemporary education oversimplifies the rational, scientific interpretation of phenomena by representing it as a pursuit of *univocal*, causal or logical dependencies. Whatever does not conform to strict determinism, whatever is associated with uncertainty, surprise, or randomness is seen as being outside of the possibility of a consistent, rational, scientific explanation.. . . The intuitions of chance and probability are influenced, and ultimately deformed, by this excessive tendency toward univocal prediction" (Fischbein, 1975; pp. 124-125).

Fischbein contends that the teaching process tends to either reject chance events as unexplicable or to impose deterministic relations where they do not exist.

From the perspective of mathematicians as well as the NCTM reform documents, conceptualizations of mathematics education should emphasize processes such as exploring patterns, inference from data, formulating conjectures, etc. It is within this conceptualization of mathematics education that chance and ambiguity enter in. Within this vein, Resnick (1988) has argued that we should not teach mathematics as a "well-structured discipline" where students practice the rules of the domain, but rather as an "ill-structured discipline" where students construct multiple strategies and viewpoints and corresponding justifications of their sensibility.

In illustration of how a classroom learning environment can reflect this kind of epistemological perspective, a collaborative activity within a primary level classroom is briefly described. The investigation was based on a suggestion in the Used Numbers' book, *Counting Ourselves and Our Family* (Stone & Russell, 1992).

A teacher challenges her combined first-second grade class to try to figure out how many children there are in a particular second grade classroom, when all the children are out of the room (and cannot be directly counted). The need for this information is framed in terms of how many cookies they should make to have enough for themselves and the other class. The children begin exploring the problem by reflecting on possible indicators of class size. For example, after one child suggests they could count pencil boxes, an informant tells her classmates that this won't work because some children share pencil boxes in this classroom. Several children think a chair count would be a good idea.

The following day, equipped with a clipboard for each pair of investigators, the children enter the empty second grade classroom. Most pairs count items that could well correspond with the

number of children in the classroom, including chairs, number of children indicated on a bar graph of birthdays by month, children's writing folders, and children's names in a poster. One pair of children count the number of Kacheena Dolls on display. They duly record the total (9), apparently oblivious of the fact that this number is a fraction of the size of any class in the school. Another pair counts a total of 13 backpacks, but immediately questions the reliability of the count as an indicator of class size, on the grounds that "lots of kids don't bring backpacks to school".

Upon return to their own classroom, the children and teacher make an ordered list of the counts they have found. The children record their counts on small squares of sticky paper (5 cm X 5 cm). The teacher constructs a bar graph, to which each pair affix their data point(s), explaining what they have counted to reach this sum. On the basis of the bar graph, the children infer there are 28 children in the other room, because 28 has the most counts.

When the children deliver the cookies, they discover that their inference was exact: the class indeed has 28 children. The teacher and a minority of the children display genuine excitement at their collaborative detective (and mathematical) work. Other children express frustration that they have failed, in that their count, the data point they contributed to the bar graph, was wrong.

Conflicting values are reflected between the teacher and the activity she has guided and many of the children new to her classroom. The teacher's task structuration and the modes of interpretation she encourages reflect an assumption that although a single count may or may not reflect the true number of children in the room, analysis of data collected across many counts of different indicators may well accurately reflect the correct number or approximate it. Furthermore, she does not lead her children to try to account for every blip in the distribution, but rather to examine the distribution for trends at the same time as they consider the issue of validity underlying outliers in the distribution. This orientation contrasts with some children's assumptions about mathematical activity; namely that, by correct application of the correct [counting] algorithm, they should have been able to get the right answer themselves. Indeed, if mathematics is simply application of the correct rule, their assumption is valid. However, in this classroom engaging in mathematics demands much more than competent algorithm-application.

Implications for assessment

Above and beyond our assessments of the learner, we need to examine how the learning environment supports or subverts the big ideas and interpretative schemas of the chance curriculum. The culture of the classroom reflects a kind of collective epistemological set, with varying degrees of convergence or conflict. Assumptions about the degree of determinism underlying phenomena or conversely the place of chance and uncertainty in the world can be reflected in the model of mathematics implicit in the curricula, the structure of problems presented to the children, practices of interpretation, and the goal-structure of classroom discourse.

IMPLICATIONS

In summary, this chapter examined assessment of students' emergent understanding and application of probability from three perspectives: (a) two key concepts underlying knowledge of the domain; (b) how beliefs come into play in whether or not students think to apply these ideas

in their attempts to make sense of patterns, and (c) how the classroom culture can support or subvert students' grasp and utilization of the big ideas underlying the statistics curriculum. The chapter also examined how teachers can assess students and classroom learning environments along these dimensions. In short, unlike most other subject areas, the teaching and learning of statistics is remarkably complex, involving both new and difficult concepts and beliefs systems that are resistant to change. Teaching of statistics can be empowered by seriously considering these various perspectives and the dynamics of their interaction.

Challenge of grasping the key constructs of randomness and chance variation. In the spirit of NCTM's recommendations that we approach the teaching of mathematics and statistics from a more conceptual perspective, identification of key concepts within the various curricular strands becomes crucial. Randomness and chance variation constitute such fundamental concepts in the context of the statistics and probability strand. The research literature reveals that these rich, multifaceted concepts can be a fruitful focus of reflection from kindergarten through high school.

Multiple dimensions of assessment: cognitive, epistemological and cultural. The goal of supporting students' appropriate interpretation of stochastic phenomena— or more specifically, their appropriate attribution of randomness and chance variation— is actually quite complex. First of all, a cognitive dimension comes into play, due to the fact that attainment of this goal presumes cognitive constructions of remarkably subtle concepts. Second, an epistemological dimension comes into play as, to a greater extent than most domains in mathematics, appropriate interpretation of stochastic phenomena also involves the students' beliefs, beliefs about the extent to which chance explains phenomena in the world. Finally, a cultural dimension comes in, as the students' beliefs about the place of chance and, more generally, stochastic versus deterministic explanations, will presumably be influenced by the values, dispositions and habits of the school and classroom culture. All three of these dimensions directly affect the success of our statistics instruction.

Assessment on the cognitive dimension. Obviously, the classroom teacher seldom has the luxury of working one-on-one with individual students to diagnose their understandings. Fortunately, small group and whole group discussions can provide a rich window onto their various understandings and beliefs. For example, returning to NCTM's question about "How many blues might we expect to get?" (NCTM, 1989, p.56), in 12 spins with a evenly-divided four sector spinner, discussions about whether or not they can predict what will happen exactly and why or why not, the source of uncertainty and variability, followed by their collecting of spinner outcome data and reflecting on how and why these data accord with and depart from their predictions, can provide a rich learning and assessment activity. In short, these crucial dimensions of students' conceptual understandings can be reflected in mathematics classes that emphasize students' explication and collaborative analysis of their emergent ideas, conjectures, reasoning processes and argumentation.

Assessment on the epistemological dimension. If we are to foster appropriate utilization of a stochastic interpretative frame, we need to also assess students' beliefs about where and how chance may enter in. Above and beyond having constructed the concept of randomness and chance variation, whether or not students will think to interpret a situation in terms of chance depends in large part on their beliefs about the place of chance and random variation in the world. For example, beliefs about the place of chance in the world may well enter into students' propensity or willingness to qualify their calculation of "9 blues" with the idea that chance is also involved here — many times the exact answer will not be 9. Furthermore, both conceptual constructs and belief systems may well affect their ability to grasp such subtle ideas as expected

value — in this case, that 3/4 is a ratio of the expected relative distributions over an infinite number of repetitions of the event, but a ratio that is only approximated across many repetitions. In short, students' beliefs about how chance fits into the world is relevant to teachers in that students' varying perspectives affect their openness to interpretative principles based on chance, as well as their inclination to use them.

Assessment of the culture of the mathematics classroom vis a vis ideas of chance. The culture of the mathematics classroom is an influential force in what students do and do not learn about ideas of chance and the place of chance in the world. In addition to assessing the students' understanding and beliefs concerning chance, teachers need to consider how the practices and values of the mathematics classroom support or hinder the ideas about chance and its application that they are intending to teach. Analysis of this implicit level of the curriculum is particularly important given the growing recognition of the power of these hidden messages and their cumulative effect on students' thinking and beliefs.