Inconsistencies in Students' Reasoning about Probability
Author(s): Clifford Konold, Alexander Pollatsek, Arnold Well, Jill Lohmeier and Abigail Lipson
Reviewed work(s):
Source: *Journal for Research in Mathematics Education*, Vol. 24, No. 5 (Nov., 1993), pp. 392-414
Published by: National Council of Teachers of Mathematics
Stable URL: http://www.jstor.org/stable/749150
Accessed: 04/03/2013 12:43

Journal for Research in Mathematics Education
1993, Vol. 24, No. 5, 392–414

INCONSISTENCIES IN STUDENTS' REASONING ABOUT PROBABILITY

CLIFFORD KONOLD, *University of Massachusetts, Amherst*
ALEXANDER POLLATSEK, *University of Massachusetts, Amherst*
ARNOLD WELL, *University of Massachusetts, Amherst*
JILL LOHMEIER, *University of Massachusetts, Amherst*
ABIGAIL LIPSON, *Harvard University*

Subjects were asked to select from among four possible sequences the "most likely" to result from flipping a coin five times. Contrary to the results of Kahneman and Tversky (1972), the majority of subjects (72%) correctly answered that the sequences are equally likely to occur. This result suggests, as does performance on similar NAEP items, that most secondary school and college-age students view successive outcomes of a random process as independent. However, in a follow-up question, subjects were also asked to select the "least likely" result. Only half the subjects who had answered correctly responded again that the sequences were equally likely; the others selected one of the sequences as least likely. This result was replicated in a second study in which 20 subjects were interviewed as they solved the same problems. One account of these logically inconsistent responses is that subjects reason about the two questions from different perspectives. When asked to select the most likely outcome, some believe they are being asked to *predict* what actually will happen, and give the answer "equally likely" to indicate that all of the sequences are *possible*. This reasoning has been described by Konold (1989) as an "outcome approach" to uncertainty. This prediction scheme does not fit questions worded in terms of the least likely result, and thus some subjects select an incompatible answer based on "representativeness" (Kahneman & Tversky, 1972). These results suggest that the percentage of secondary school students who understand the concept of independence is much lower than the latest NAEP results would lead us to believe and, more generally, point to the difficulty of assessing conceptual understanding with multiple-choice items.

Probabilistic reasoning is receiving increased attention among mathematics educators and researchers, due largely to a new resolve to transform precollege mathematics education in the United States. Fortunately, a considerable body of research in this area has been amassed over the past few decades by psychologists and decision theorists who have long been interested in how people reason about probabilities when they make decisions. What does this research tell us about the probabilistic reasoning of the typical adult in this country? According to Slovic, Fischoff, and Lichtenstein (1976), most people do not reason in accord with accepted theory; they "systematically violate principles of rational decision-making."

This research was funded by grant MDR-8954626 from the National Science Foundation to Clifford Konold. The opinions expressed here do not necessarily reflect those of the Foundation. Portions of this research were presented at the eleventh annual meeting of the North American Chapter, International Group for the Psychology of Mathematics Education, Rutgers University, September 1989, and at the Third International Conference on Teaching Statistics, Dunedin, New Zealand, August 1990. We are grateful to Joan Garfield for the data from courses at the University of Minnesota and to Ruma Falk and Amy Robinson for their comments on earlier drafts.

Conflicting claims have been made both about people's understanding of basic probabilistic and statistical concepts and about the ease with which these concepts can be learned. A large body of research indicates that people employ a small set of heuristics when making probability judgments (e.g., Tversky & Kahneman, 1974; Shaughnessy, 1992). These heuristics, one of which we describe later, often result in quick and generally reasonable judgments but can lead to judgments that are strongly at odds with probability theory. Additionally, research by Konold (1989) suggests that some college undergraduates reason about uncertain outcomes using a fundamentally nonprobabilistic "outcome approach." In contrast to these findings are claims of Piaget and Inhelder (1975) that by the age of 12, most children acquire basic probability concepts even without formal instruction. The latter conclusion seems to be supported by successful performance on probability problems included in the fourth National Assessment of Educational Progress (NAEP). In addition, Nisbett and his associates (e.g., Fong, Krantz, & Nisbett, 1986) claim that basic probability concepts, such as the law of large numbers, can be taught with reasonable success to undergraduates in roughly half an hour.

Our position is that, in some sense, both of these general claims are correct. That is, the typical person has knowledge about a variety of uncertain situations, but that knowledge is incomplete and not integrated. Different problems access different pieces of this knowledge. Thus in one problem, a person may appear to reason correctly, but in another, this same person may reason in ways that are at variance with probabilistic and statistical theory. Other researchers have noted and tried to explain inconsistencies across different problem types. For example, Nisbett, Krantz, Jepson, and Kunda (1983) have found good performance on problems that (a) involve repeatable processes with a finite set of symmetric outcomes (e.g., rolling a die), (b) comprise outcomes produced by a method associated with chance (e.g., blindly drawing from a set of well-mixed objects), and (c) are widely recognized as being unpredictable and capricious (e.g., the weather). As features of a statistical problem begin to deviate from these prototypes, people revert to nonstatistical and inappropriate ways of reasoning.

We propose that incorrect reasoning frequently occurs even with prototypical chance events and that a subject can switch from correct to incorrect reasoning while reasoning about what an expert would consider to be the same situation. To account for these types of inconsistencies, it is critical to understand the beliefs and reasoning processes that underlie the various answers that subjects give. Consider, for example, the following problem:

> A die is painted white on one side and black on the other five sides. If the painted die is rolled six times, which of the following two outcomes is most likely?
>
> *a*) Black side up on five rolls and white side up on the other roll.
>
> *b*) Black side up on all six rolls.

According to the binomial calculation, *a* is more likely, with a probability of 0.402 compared to 0.335 for *b*. In an interview study by Konold (1989), 5 of 12 undergraduates answered this problem correctly. We cannot conclude, however, that subjects who gave the correct answer were reasoning normatively (i.e., according to accepted principles of probability theory), because none of them computed the probabilities as above. Their answers may have been based on the "representativeness heuristic" (Kahneman & Tversky, 1972) according to which the judgment of how likely it is that a particular sample was drawn from a population is made by considering how *similar* the sample is to the population. One task Tversky and Kahneman (1974) have used to investigate this heuristic involves guessing the occupation of a person based on brief descriptions like the following:

> Steven is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail. (p. 1124)

Given the choice of Steven's being a mathematician or a lawyer, people using the representativeness heuristic would make that judgment by comparing the description of Steven to their stereotypes of people in those occupations. If the description fit better the stereotype of a mathematician, then Steven would be judged as more likely to be a mathematician than a lawyer. Although this method is not irrelevant to making probability judgments, it is flawed because it does not take into account factors other than similarity that affect probability. For example, the fact that there are many more lawyers than mathematicians in this country increases the probability of Steven's being a lawyer relative to that of his being a mathematician.

In the case of the painted-die problem above, people using the representativeness heuristic would answer correctly, since the ratio of elements in the most likely outcome *a* is identical to the ratio of elements in the population distribution. Thus, sample *a* is more similar to the population than sample *b* (see Kahneman & Tversky, 1973, for discussion of a similar problem). The important point is that correct performance on this item cannot be used to discriminate between those reasoning formally via probability theory and those using the representativeness heuristic.

What was the logic used by the seven subjects who thought *b* (black side up on all six rolls) was more likely? One theory that accounts for this response is the "outcome approach" (Konold, 1989; 1991b). Given an uncertain situation, people using the outcome approach do not see their goal as specifying probabilities that reflect the distribution of occurrences in a sample but as *predicting* the results of a *single* trial. Those applying the outcome approach to the painted-die problem derive an answer by thinking about the outcome of each of the individual trials, as suggested in the following subject transcript from Konold (1989, p. 84):

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

This subject considers each trial separately, predicting on each trial that a black will come up. Considering the six predictions together leads to the conclusion that six blacks is the more likely outcome. This reasoning is consistent with the outcome approach because it involves making yes-or-no predictions of the results of individual trials.

Although the outcome approach is only one possible explanation for incorrect answers to the painted-die problem, there is evidence that it captures the reasoning of some subjects. Konold (1989) found that a measure of adherence to the outcome approach generated from an analysis of subjects' answers to a different set of problems predicted responses to the painted-die problem. Konold (1989) also found that subjects' responses were not consistent across problems. Many subjects who appeared to reason according to the outcome approach on one problem seemed to reason correctly on another and showed evidence of using a heuristic approach on yet a third. Different problems appeared to induce subjects to retrieve different (and perhaps incompatible) beliefs.

To summarize, in the painted-die problem, subjects chose one of two alternatives. If they chose the correct answer, they may have been reasoning either normatively, as an expert might, or according to some heuristic, such as representativeness. If they chose the incorrect answer, they certainly were not reasoning normatively and may have been reasoning according to the outcome approach. Furthermore, given a different problem, some subjects appeared to switch from normative reasoning to a heuristic or outcome approach, and vice versa.

The two studies reported here were designed to investigate subjects' probabilistic reasoning. In Study 1, we investigated the possibility that correct answers to probabilistic questions may be based on nonnormative reasoning. In Study 2, we examined within-subject inconsistencies in reasoning.

STUDY 1

The following HT-sequence problem is similar to one used by Kahneman and Tverksy (1972) to distinguish those who reason normatively about probability from those who apply the representativeness heuristic:

> Which of the following is the most likely result of five flips of a fair coin?
>
> *a*) HHHTT
> *b*) THHTH
> *c*) THTTT
> *d*) HTHTH
> *e*) All four sequences are equally likely.

The correct answer is *e,* that the sequences are equally likely. Indeed, each of the 32 possible ordered sequences has probability $(1/2)^5$. As with the painted-die problem, it is not necessary that people know how to calculate probabilities in order to answer this problem correctly. If they viewed the outcome of each flip as *independent* of the outcomes of all prior flips, they could deduce that all sequences are equally likely without computing the probability. We would expect that subjects employing the representativeness heuristic would choose THHTH as being more likely than either THTTT or HTHTH. There are two ways in which THHTH might be considered more representative of coin flipping (cf. Tversky & Kahneman, 1974). First, in contrast to THTTT, it has a nearly equal number of heads and tails, better reflecting the fact that heads and tails are equally likely. Second, in contrast to HTHTH, the sequence of THHTH is more mixed up and hence representative of the randomness of coin flipping.

Not all subjects give such a heuristic-based answer. In fact, when we collected pilot data on this coin problem, we discovered that the majority of subjects selected the correct answer. However, we also asked subjects to provide a written justification for their answer. Many subjects who answered correctly gave justifications that suggested they were reasoning according to the outcome approach. Many of them, for example, said that it was impossible to predict which outcome would occur. This response suggests they believed they were being asked what *actually* would happen if a coin were flipped five times. These results are quite different from those reported by Kahneman and Tversky (1972). Of the 92 subjects answering their problem, which dealt with the order of birth of girls and boys in a family, 82% responded in a manner consistent with the representativeness heuristic, judging the sequence BGBBBB to be less likely than GBGBBG. In their problem, only these two possible sequences were offered for comparison, and subjects were asked to estimate the frequency of occurrence of one sequence given the frequency of the other. As we argue in this paper, what may seem like trivial differences in the presentation of two problems can result in radically different patterns of responses.

To test whether some correct answers in the HT-sequence problem were based on nonnormative reasoning, we devised a variant of the problem. In this version, the follow-up question "Which of the above sequences would be least likely to occur?" was placed immediately below the original question. For reasons that we will discuss later, we expected that those reasoning according to the outcome approach on the first part of the problem would adopt a different approach and choose a different answer on the second part.

Method

Materials and Procedure

The two items below were included in questionnaires along with other items on probability and statistics.

Four-heads problem. A fair coin is flipped four times, each time landing with heads up. What is the most likely outcome if the coin is flipped a fifth time?

a) Another heads is more likely than a tails.
b) A tails is more likely than another heads.
c) The outcomes (heads and tails) are equally likely.

HT-sequence problem.

Part 1. Which of the following is the most likely result of five flips of a fair coin?

a) HHHTT
b) THHTH
c) THTTT
d) HTHTH
e) All four sequences are equally likely.

Part 2. Which of the above sequences would be least likely to occur?

Each item appeared on a separate page, and subjects were instructed not to go back to a page once turned. Parts 1 and 2 of the HT-sequence problem appeared on the same page.

Subjects

Summermath. Both items were administered as part of a nine-item pretest to 16 high school students on the first day of a workshop on probability. This workshop was offered in 1987 as part of Summermath, a 6-week residential program for young women sponsored by Mount Holyoke College. Summermath recruits nationwide; its participants represent a broad range of mathematical ability.

Remedial mathematics. Twenty-five undergraduate students enrolled in the Spring 1987 semester of a remedial-level mathematics course at the University of Massachusetts volunteered to participate in a study on probabilistic reasoning. Probability was not a topic covered in this course. The four-heads and HT-sequence problems were among 14 items completed.

Statistical methods. Both items were administered in the fall of 1987 to 47 students as part of a 10-item precourse survey for a statistical methods course in the College of Education at the University of Minnesota. This course is the first of a three-semester methods sequence required of all advanced-degree candidates in psychology and education.

Results and Discussion

Four-Heads Problem

Overall, 86% of the subjects correctly chose option *c* for the four-heads problem. Not surprisingly, the performance of the remedial students was the poorest (70% correct) and that of the statistical-methods students the best (96% correct). The most popular alternative answer was the one consistent with the so-called gambler's fallacy, that an outcome of tails is more likely after a run of heads. This option was selected by 22% of the Remedial students, 19% of the Summermath students, and 4% of the Statistical-methods students.

Green (1982) included a similar item on a probability-concepts test that was administered to over 4,000 English school pupils of ages 11–16. The item differed slightly in that the coin was said to have previously landed heads up five times (rather than four), and a fourth option, "Don't know," was included. Overall, 75% of the students answered correctly, with 67% of the 11–12-year-olds answering correctly compared to 80% of the 15–16-year-olds. However, unlike our results, subjects in Green's study chose another heads (11%) about as often as they chose tails (12%) on the last flip.

From our results and those of Green, one might conclude that by the age of 12, only a small percentage of people reason according to the gambler's fallacy and that the majority believe in the independence of trials in coin flipping.

HT-Sequence Problem

Performance on the HT-sequence problem is summarized in Table 1. The majority of subjects (72% overall) correctly chose option *e*. As with the four-heads problem, performance on the HT-sequence problem suggests that by late adolescence, people believe that successive trials in coin flipping are independent and that therefore all possible sequences of flipping a fair coin five times are equally likely. Contrary to the findings of Kahneman and Tversky (1972), only a small percentage of people seem to use the representativeness heuristic in reasoning about these types of problems.

Table 1
Percentage of Responses Identifying Each Sequence as the Most and Least Likely Outcome of Five Flips of a Coin (Study 1)

	Group						Total	
	Remedial		Summermath		Stat. methods			
Prob. version/ sequence	Most	Least	Most	Least	Most	Least	Most	Least
a) HHHTT	17.4	8.7	0	6.7	0	7.3	4.7	8.0
b) THHTH	13.0	4.3	25.0	0	2.1	2.4	9.3	2.5
c) THTTT	4.3	8.7	0	33.4	2.1	26.8	2.3	22.8
d) HTHTH	0	43.4	6.3	40.0	10.6	17.1	7.0	29.1
e) Equal	60.8	34.8	68.8	20.0	78.7	46.3	72.1	38.0
f) *a, b, d**	4.3	0	0	0	6.4	0	4.7	0
(N)†	(23)	(23)	(16)	(15)	(47)	(41)	(86)	(79)

*Response added by subjects who indicated that options *a, b,* and *d* were equally likely and that option *c* was least likely to occur.

†The sample sizes, N, are not always equal for the least and most likely versions of the problem because some subjects left Part 2 blank.

However, performance on the follow-up question suggests that the majority of our subjects were not reasoning correctly. The result of particular interest is the percentage of correct responses to the question of which sequence is *least* likely. Overall, only 38% of the subjects responded that all four sequences were equally (un)likely. In other words, roughly half the subjects who answered Part 1 correctly responded again that the sequences were equally likely; the others selected one of the sequences as least likely. Hereafter, we will refer to this pattern of response as the M-L (most-least) switch. It thus appears that many subjects who chose the correct answer in Part 1 nevertheless do not believe that all sequences are equally likely.

An Explanation Based on the Outcome Approach

One account of these logically inconsistent responses is that subjects reason about the two parts of the question from different perspectives. When asked about the most likely outcome, some believe they are being asked to *predict* what will happen and give the answer "equally likely" to indicate that all the sequences are *possible*. This reasoning is consistent with the "outcome approach" to uncertainty (Konold, 1989). As mentioned in the introduction, outcome-oriented individuals, when asked the probability of an event, typically interpret the request as one to specify what *will* happen. Rather than interpreting the four-heads and HT-sequence problems as inquiries concerning the probability of various outcomes, they think they are being asked what *will* happen on the fifth trial, or which five-character sequence *will* occur, respectively. Since the 50% probability associated with coin flipping suggests to them that no prediction can be made, they choose the answer "equally likely." In this context, *equally likely* does not mean that the sequences have the same numeric probability of occurrence, but that there is no basis for making a prediction of what will happen.

In addition to choosing an option, students were asked to give a brief justification for their answer to each problem. Many justifications offered little insight into the underlying rationales. Three justifications that were informative are included in Table 2. All three subjects selected answer *e* (equally likely) in response to the question of which sequence was most likely. The answers each of these subjects selected in response to the question of which sequence was least likely are noted in brackets next to their rationales in Table 2. The small number of these responses (3) constitutes rather limited support for our hypothesis that the M-L switch is accounted for by the outcome approach. We included them here, however, to illustrate the types of rationales we regard as consistent with this hypothesis. These justifications indicate that in answering the question of which sequence is most likely these subjects focused more on the nonpredictability of coin flipping than on the probability of the various sequences. In their rationale for the most-likely sequence, all three subjects mentioned that any of the sequences *could* occur. An implication of this justification is that because one cannot rule out the occurrence of any of the sequences, one cannot predict with assurance which sequence will occur. This reasoning is consistent with the outcome approach in that it is based on the presumption that the objective in this situation is to predict what will occur.

Table 2
Reasoning in the M-L Switch. Written Justifications of Three Subjects Who Responded Inconsistently to Parts 1 and 2 of the HT-Sequence Problem (Study 1)

	Version of problem	
S#	Part 1: Most likely	Part 2: Least likely
S_{15}*	[*e*] Anything can happen with probability. The chances of some of the examples are least likely to occur (*a,c*) but it can happen.	[*a, c*] These chances are least likely to occur because they happen the same side in a row.
S_{16}	[*e*] They all could occur.	[*c*] Because it is least likely to occur when you have almost a perfect score.
R_2	[*e*] It's a chance game. No skill is involved, therefore all could likely occur by chance.	[*a*] Receiving 3 heads in a row seems unlikely, but could very well occur.

*S = Summermath. R = Remedial.

The answers and accompanying justifications for the question of the least likely sequence seem consistent with the representativeness heuristic. Note (in Table 1) that the majority of subjects selected HTHTH or THTTT as the least likely to occur. As argued earlier, the former sequence is unrepresentative because it appears too ordered, and the latter sequence has too many tails. The excerpts in Table 2 all focus on the improbability of runs in the chosen sequence. Long runs violate the representativeness heuristic because they (*a*) appear nonrandom and (*b*) produce an excess of one outcome over the other.

A paraphrasing of the thinking that might be responsible for the M-L switch is, "I can't say which sequence will occur [most likely version], but I think sequence *x* is particularly unlikely [least likely version]." If subjects who gave inconsistent answers were reasoning according to the outcome approach on the first part of the problem, why did they switch to answering the second part of the problem on the basis of the representativeness heuristic? One possibility is that although the prediction scheme of the outcome approach fits the question when worded in terms of the most likely result, it ceases to fit the question when the wording is changed to ask about the least likely result. This is because predictions involving more than two mutually exclusive events are typically made about the occurrence rather than the nonoccurrence of events. Thus it seems natural to predict which of ten horses will win a race but would strike many as strange to predict which horse would not win because nine of them will not win. This asymmetry between the two parts of the problem from the point of view of a prediction scheme may thus produce for some subjects a change in perspective, from an outcome approach to the more probabilistically-oriented representativeness heuristic. Because their answers are based in different frameworks, subjects may not perceive any contradiction between them.

STUDY 2

In Study 2, we interviewed subjects as they reasoned about the same two problems. The major objective of these interviews was to explore the reasoning of subjects who committed the M-L switch, and for this purpose we included several follow-up questions. The interviews revealed additional instances of the M-L switch as well as other inconsistencies that we had not anticipated. To account for these, we suggest that subjects not only switch among incompatible perspectives of uncertainty but reason at times from basic beliefs they hold about coin flipping. For example, the same subject may justify one answer by stating that a coin is unpredictable, and another answer by claiming that certain outcomes of coin flipping are more likely than others. Logically, these beliefs are not contradictory; they are, however, incomplete. Contradictory statements (and statements at variance with probability theory) might occur when subjects apply these beliefs beyond their appropriate domain.

Method

Twenty subjects (12 women and 8 men) were recruited from undergraduate psychology courses at the University of Massachusetts. Eleven of the subjects had taken, or were currently enrolled in, a course in which statistics had been taught. The subjects participated in an hour-long videotaped interview that included several other questions concerning various aspects of probability. The four-heads problem was presented as in Study 1. The HT-sequence problem was modified slightly as shown below:

Part 1. Which of the following sequences is most likely to result from flipping a fair coin five times?

a) HHHTT
b) THHTH
c) THTTT
d) HTHTH
e) All four sequences are equally likely.

Part 2. Which of the following sequences is the least likely to result from flipping a fair coin five times?

a) HHHTT
b) THHTH
c) THTTT
d) HTHTH
e) All four sequences are equally unlikely.

To stress the difference between the two parts of the question, the words *most* and *least* were underlined. Also, the options were now repeated in Part 2 of the problem, with the option *e* reworded to match the wording of the item stem in terms of the *un*likelihood of the sequences. The two parts of the problem were again presented on the same page.

A list of the procedures including the planned probes is shown in Table 3.

Table 3
Interview Procedures (Study 2)

1. Subject reads aloud and answers the four-heads problem.
2. Interviewer asks: "Can you tell me why you think that [repeat wording of selected option]?"
3. Subject reads aloud and answers Part 1 of HT-sequence problem.
4. Interviewer asks: "Why did you answer [name selected option]?"
5. Subject reads aloud and answers Part 2 of HT-sequence problem.
6. Interviewer asks: "Why did you answer [name selected option]?"
7. Interviewer: "There are two ways to interpret the outcome HHHTT, for example. One way is to think of flipping a coin five times and getting three heads and two tails, in any order. The other way to interpret HHHTT is to think about flipping the coin five times and getting H on the first three flips and T on the last two — in other words, getting HHHTT in exactly that order."
   a) "In this problem, how were you interpreting these outcomes?"
   b) "Were you thinking about getting heads and tails in the exact order as listed, or in any order?"
   c) If the subject says she or he was not paying attention to order, the interviewer asks the subject to do the problem over, this time considering the exact order of the sequences.
8. If the subject selected an option (x) as least likely in Part 2, the interviewer reminds the subject of the sequences he or she selected as least likely and does the following:
   a) Asks: "What is the probability of x occurring?"
   b) Interviewer selects two of the options that were not selected and asks for each: "What is the probability of option y occurring?"
9. If a subject selected option *e* in Part 2, the interviewer asks:
   a) "What is the probability of *d* occurring?"
   b) "What is the probability of *b* occurring?"

*10. Subject is again given the four-heads problem. Interviewer asks:
   a) "What is the probability of another heads occurring?"
   b) "What is the probability of getting a tails?"

*We waited until the end of the interview to ask for probabilities in the four-heads problem because we feared that the question might induce some conflict and that if asked early could influence responses to the HT-sequence problem.

Results and Discussion

Table 4 codes correct (+) and incorrect (–) responses for each subject on six aspects of the two questions. The percentages of correct responses to each question are listed along the bottom row of the table. There were no significant differences in the mean number of correct responses based on either gender or prior statistics instruction.

Of primary interest in Study 2 was the consistency of a subject's responses over the two parts of the HT-sequence problem. Fourteen (or 70%) of the subjects responded correctly in Part 1 that all sequences are equally likely. In spite of changes to the HT-sequence problem, which were intended to make inconsistencies more apparent to the subjects, four of the subjects (S_3, S_5, S_{11}, and S_{14}) committed the M-L switch, as indicated by a plus in column 3 (Most) and a minus in column 4 (Least).

Table 4
Correct and Incorrect Responses to the Four-Heads and HT-Sequence Problems (Study 2)

		Problem					
		Four heads		HT-sequence			
	S#	Equal	*P* = .5	Most	Least	Equal *P*s	Sum < 1
	6*	+	+	+	+	+	+
A	16	+	+	+	+	+	+
	20	+	+	+	+	+	+
	13	+		+	+	+	+
	17	+	+	+	+	n	n
B	2*	+		+	+		
	15	+	+	n	n	n	n
	12*	+	+	+	+	+	–
	14*	+	+	+	–	+	+
	19	+	+	+	+	–	+
	8*	+		+	+	–	+
	4*	+	+	–	–	+	+
C	5	+	+	+	–	+	–
	11*	+	+	+	–	–	+
	18	+	+	+	+	–	–
	9*	+		–	–	+	+
	3	–	+	+	–	–	–
	7*	+	+	–	–	–	–
	1*	–	+	–	–	–	–
D	10*	–	–	–	–	–	+
% Correct		85	94	74	53	53	65

Key: [+] correct response; [–] incorrect response; [n] responded that it was not possible to answer; [ ] missing values indicate that the question was not asked; [*] prior or concurrent course(s) in probability and statistics.

Row Headings: [A] subjects answered all questions correctly; [B] subjects showed no obvious inconsistencies but protocols are incomplete; [C] subjects gave inconsistent responses; [D] subject gave consistent but nonnormative responses.

Column Headings: [S#] subject number; [Equal] heads and tails are equally likely outcomes in Four-Heads problem; [*P* = .5] probability of heads (or tails) is .5 in four-heads problem; [Most] all sequences are equally likely in Part 1 of HT-sequence problem; [Least] all sequences are equally likely in Part 2 of HT-sequence problem; [Equal *P*s] HT sequences are assigned equal probabilities; [Sum < 1] sum of probabilities given for mutually exclusive sequences is less than 1.

Other Inconsistencies

The additional probes included in the interview revealed a number of other inconsistencies. To convey an overall impression of how consistent subjects were in their responses, subjects are ordered in Table 4 according to the total number of correct responses. The 4 subjects listed in the top group, A, of the table answered all the problems correctly. The subject in the bottom group, D, answered consistently in accord with the representativeness heuristic. The 3 subjects in group B showed no obvious inconsistencies but said it was not possible to answer one or more questions, as indicated by an *n* in Table 4. The 12 subjects in group C showed various inconsistencies in their responses and thus are of greatest interest here. The most salient of these inconsistencies are described below.

Qualitative answers versus probabilities: HT-sequence problem. Subjects who respond that the sequences are equally likely (and unlikely) ought then to give equal probabilities to the options. However, S_8, S_{18}, and S_{19} each responded correctly to the most and least likely versions of the HT-sequence problem and then assigned unequal probabilities to the sequences. Thus, it appears that even some of those who responded correctly to *both* versions of the HT-sequence problem do not believe the sequences have equal probabilities of occurrence.

In contrast, some subjects who indicated a belief that one of the sequences was more (or less) likely than the others assigned equal probabilities to the sequences. S_4 and S_9 selected option *c* as least likely, but then went on to assign the same probability to option *c* as to the other options (10–20% in the case of S_4; 20% in the case of S_9).

Qualitative answers versus probabilities: Four-heads problem. Subjects reasoning normatively about the four-heads problem should respond that heads and tails are equally likely outcomes and assign equal probabilities to each outcome. Although most subjects responded in this manner, S_3 and S_1 selected tails as the more likely outcome and then assigned both heads and tails an equal probability of 50%. S_{10} was the only other subject to select tails as more likely and the only one to give a higher probability to tails. Thus his reasoning, though incorrect, was consistent.

Responses on the four-heads versus HT-sequence problems. Subjects correctly answering the four-heads problem would seem to be exhibiting an understanding of the independence of successive trials in coin flipping. Given this understanding, these subjects ought to regard the various sequences in the HT-sequence problem as equally likely. S_4, S_7, and S_9 all gave correct answers to the four-heads problem and incorrect responses to both versions of the HT-sequence problem.

Constraint on the sum. Given that the sum of the probabilities of all 32 sequences is 1, the sum of the probabilities of a proper subset of these sequences should be less than 1. Six subjects gave probability values whose sum equaled or exceeded 1. It should be noted that because subjects were asked for the probabilities of only 2 or 3 of the sequences, these results provide a conservative estimate of the number of subjects who do not realize that the sum of the probabilities of mutually exclusive events cannot exceed 1.

Possible Explanations for Inconsistent Responses

In this section, we present two different accounts of the various inconsistencies revealed in Studies 1 and 2. We suggest that some inconsistencies, such as the M-L switch, result from the application of different conceptual frameworks that people use to think about a large range of uncertain situations. Other inconsistencies, in contrast, seem to result from the application of maximlike beliefs that may apply only to specific situations, such as coin flipping.

Availability of multiple frameworks. In the introduction, we mentioned distinctions among what we consider to be three general frameworks for making probability judgments: the normative, formalized framework used by experts to compute probabilities; the informal judgment heuristics used in everyday situations to arrive at quick assessments of probabilities; and the single-trial-prediction framework of the outcome approach. The typical adult probably reasons according to each of these frameworks at various times. Inconsistencies would result, however, if a subject switched among these frameworks in thinking about different aspects of the same situation. The suggestion that people have multiple perspectives or frameworks about probability is not new (e.g., Nisbett, Krantz, Jepson, & Kunda, 1983; Tversky & Kahneman, 1981). Indeed, Tversky and Kahneman (1971) demonstrated that experts well versed in normative thinking can be "fooled" into using the representativeness heuristic with problems that are sufficiently complex. We suggest, however, that different frameworks can be employed almost simultaneously in reasoning about the same situation.

In the discussion of Study 1, we speculated that the M-L switch resulted from a change in perspective from an outcome approach in Part 1 of the problem to a representativeness heuristic in Part 2. We suggested that statements such as "anything could happen," which were used by a few subjects to justify the response of "equally likely," were indicative of the outcome approach. The interview format gave us further opportunities to evaluate statements indicative of this interpretation.

Table 5 is a listing of rationales given in Study 2 by subjects who selected the correct option in Part 1 of the HT-sequence problem. Seven of the rationales for the equally likely response (indicated in the table with asterisks and boldface) were judged to involve arguments consistent with the outcome approach, for example, that anything was possible or that it was not possible to know what would happen. Of these seven subjects, two (S_3 and S_5) gave an inconsistent answer to the question of which sequence was least likely, and one (S_8) gave nonequal probabilities for the various sequences.

One of the more revealing statements was made by S_{18}. She gave consistent answers to the most and least likely versions of the HT-sequence problem, but when asked to give a value for the probability of sequence *d* (HTHTH), she responded,

Table 5
Verbal Justifications of Subjects Who Answered (e) Equally Likely to Part 1 of the HT-Sequence Problem (Study 2)

S#	Version of problem: Part 1: Most likely	Part 2: Least likely
2*	[*e*] It's an equal chance. So each time you flip a coin, it can come up either heads or tails, and they can do it in all different arrangements. **So anything is really likely.**	[*e*] They each have the same chance. It's not most likely or least likely, I don't know.
3*	[*e*] I just think that they—**any of them could happen,** cause it's not like all heads or all tails.	[*c*] Because tails would have to come up the last three times.
5*	[*e*] You can't tell. It's a game of chance. 50/50—**it could be anything.**	[*c*] Well, you can't really say. But if I had to say one, I guess I would say *c* because it's more tails than heads, but you can't really say that though.
6	[*e*] I guess cause each time you flip a coin it's 50/50 so, I don't know. Because the probability of the combination of all the sequences are all the same.	[*e*] Because the combination, the probability of the combination of those are all the same.
8*	[*e*] They are all equally likely. I've just never had that [HTHTH] happen before.... But **just because it hasn't happened to me doesn't mean that it doesn't happen,** I guess. So they're probably all equally likely.	[*e*] Because a coin doesn't have memory of where it landed on the first time, you know, so it doesn't switch off, say "Well, I was heads last time, so I'm going to be tails next time."
11	[*e*] In each sequence there's either—heads and tails has just one more. And they're equally likely because heads and tails are equally likely on each outcome.	[*c*] Because it has tails coming up four times, whereas the other ones have about the same amount.
12*	[*e*] **Because it could come out anything.** Every time you flip it it's a 50/50 chance—**it could come out either way.** It could come out all heads or all tails, either way.	[*e*] Again, because it could come out any way. I don't think it makes any difference. It could come out any way.
13	[*e*] Because you can't tell. It's totally random.	[*e*] Because when you flip a coin it has an equal chance of being either. And it could just as easily land heads every time, or not.
14	[*e*] Every time a coin is flipped there's a 50% chance of it occurring either way.	[*a*] Just because to me it appears a little bit more predetermined as compared to random, where if you just flipped a coin five times, it should be completely random, the outcomes.
16	[*e*] Each time you flip there's a 50/50 chance, and so maybe *a* would be exactly reversed, maybe *d* would be exactly reversed. I don't think that there is a single most likely way for it to come out.	[*e*] Again, I feel the same way.
17	[*e*] Because there is no way you can—You know, any sequence—If you flip a coin forever, it will be half heads and half tails.	[*e*] For the same reason.
18	[*e*] All are equally likely because it's either heads or tails.... The probability is in the fact that it's heads or tails, not in the number of times you flip it.	[*e*] Because once again, it's just the heads and tails, not the [length of the] sequence.
19*	[*e*] **You can't really predict it.**	[*e*] You can't choose which side it's going to land on.
20*	[*e*] **There is no way to know what's going to happen.** There are no variables. It's just heads or tails. 50% chance on each side.	[*e*] Same thing.

*Response coded as indicative of outcome approach based on the statement in boldface.

> I'd say that it's very probable. I'd say that there's a 50/50 chance that you would get it that way, that—I guess I'm just—because I just look at flipping a coin as either a yes-or-no kind of thing. So that, "sure that's possible," and "sure that's possible," and "sure they're all possible." But maybe probability is a different word in terms of possible. But I would say, though, what I'm considering as probability, that that's, you know, very probable.

This subject seems to be aware that her interpretation of probability in terms of the *possible* may not be standard. But reasoning according to her interpretation, she gives sequence *d* a 50% probability, meaning, to her, that the sequence is *possible.* But when asked the probability of sequence *b* (THHTH), she adapted her interpretation to allow different degrees of *possibility.*

> S_{18}: Maybe there's something about this perfect order of HTHTH that sort of says—is not that probable, because it's so perfectly in order. Whereas *b* is a little more askew, and so maybe that seems even more probable to me.
>
> *I:* So, if you had to put a number, like a percentage, for probability on sequence *d* or on sequence *b,* what would that be?
>
> *S:* I'd say 50% for *d* and, maybe, like 65–70% for *b.*

She may have preferred to keep *d* at 50% in order to maintain its meaning of *possibility,* allowing values greater than 50% to indicate higher degrees of possibility. Gauging possibility from the benchmark value of 50% led her to give percentages for mutually exclusive events that exceeded 100%. But she is clearly not using percentages as probabilities and thus is not bound by this constraint.

The use of multiple frameworks and midstream framework switching may also account for inconsistencies between the alternatives subjects selected and the probabilities they later assigned. In the HT-sequence problem, for example, some subjects may respond that the sequences are equally likely (in accord with the outcome approach) and not think it inconsistent to assign higher probabilities to some of the outcomes than to others by virtue of the representativeness heuristic. The question in Part 1, "Which sequence is most likely?" may be interpreted as a request to make a prediction, whereas the task of assigning probabilities is not. It could be that subjects' subjective probabilities for the sequences, although not equal, are not different enough, in their judgment, to provide a basis for predicting one sequence over another.

Conflict among maxims concerning coin flipping. We found it difficult to account for many of the inconsistencies apparent in Table 4 in terms of the framework-switching hypothesis. An analysis of the interview protocols suggested that some of these inconsistencies resulted from applying maxim-like beliefs about coin flipping. Below is a list of assertions that seemed to serve as maxims that some subjects applied to answer the various questions about coin flipping:

- One cannot predict for certain the results of coin flipping.
- The outcomes of repeated trials vary unsystematically between heads and tails.

- A coin has no memory.
- Heads and tails are equally likely.
- Heads and tails occur about equally often in a sample of flips.

Statements like these are evident in Table 5 and were made repeatedly by subjects during the interview. These statements were offered by students as self-evident and were used to justify other claims.

Nisbett and Ross (1980) suggested that in many situations, people reason from culturally shared beliefs as captured in maxims, epigrams, anecdotes, fables, and so on. For example, many people use the maxim "absence makes the heart grow fonder" to explain or predict the result of a long separation between lovers. One of the difficulties encountered in reasoning from such beliefs is that they do not contain information about when and how they should be applied. Although they seem to capture the collective experience of a culture, they nevertheless lack the specificity required for rational decision making. Indeed, examples of contradictory maxims abound, for instance, "out of sight, out of mind." Because these beliefs come with little or no information concerning boundary conditions, selection of a particular maxim seems a matter of caprice, and thus people may easily reason in contradictory fashion given situations that would appear to be nearly identical.

Similarly, subjects in our study who applied the maximlike assertions listed above may not have a firm grasp on the underlying normative framework in which these assertions are properly qualified. When taken separately, these assertions can be misleading. For example, the unqualified belief that heads and tails occur about equally often in a sample of flips might lead a person to conclude that tails must be more likely than heads after a long run of heads. However, this belief would seem to contradict the belief that the coin has no memory. The belief that sequences of flips vary unsystematically between heads and tails may suggest that a *scrambled* sequence is more likely than an orderly sequence. However, this belief contradicts the belief that the outcomes of coin flipping are unpredictable.

According to this account, various aspects of a problem may invoke particular beliefs, or maxims, which then serve as a premise on which to base an answer. Slight modifications to the problem alter the beliefs that are invoked. Inconsistencies occur as a result of reasoning from incompatible assertions. For example, a run of four heads may not seem long enough to abandon the expectation that with a fair coin, heads and tails are equally likely. Thus, the person responds that "it's still 50/50 whether you get a heads or a tails." But given a longer run of, say, six heads, the deviation from the expected even split of heads and tails may be too large, so that the person reasons, on the basis of equality in samples, that tails is now more likely. At some point, the belief in equality in samples may come to dominate the belief that the probabilities are 50/50. Although subjects may not typically be aware that their various beliefs support different expectations, there is always a potential for conflict.

Such a conflict between two strongly held beliefs is evident in the written justification given by a statistical-methods subject to the HT-sequence problem in Study 1:

> I come up with three different thoughts, and can't decide which is accurate. THTTT means that four out of five times you got T. 50/50 would be closer to [a split of] 3/2. But at the same time, each flip is 50/50.

For this student, the conflict is between the belief that (*a*) samples ought to have nearly equal numbers of heads and tails, which would make THTTT less likely than, say, THHTH, and (*b*) each flip (and therefore perhaps each sequence) is equally likely.

The inconsistent responses by S_4, S_7, and S_9 to the four-heads and HT-sequence problems (see Table 4) may be indicative of reasoning from these maximlike beliefs. The four-heads problem asks about the probability of heads versus tails conditioned on the specific result of four previous flips. The primary beliefs that this problem seems to invoke are that (*a*) coin flipping is unpredictable, (*b*) the coin has no memory, and (*c*) the two outcomes are equally likely. There are at least two possible reasons for these beliefs to be cued. One is that coin flipping is strongly associated with the phrase "50/50," which serves for some as a synonym for "unpredictable." If this was the only reason, however, then one would expect similar beliefs to be cued in the HT-sequence problem, since it too involves a coin with a 50/50 chance. However, in the four-heads problem, the fact that the coin is about to be flipped once may be more salient to subjects than the two sequences that could result: HHHHT and HHHHH. Accordingly, the belief in unpredictability, or equal likelihood, is cued rather than the belief that samples ought to contain roughly equal numbers of heads and tails. In the HT-sequence problem, however, the sequences are more salient than single flips. As a result, beliefs in the equality of frequencies of heads and tails in samples and the irregularity of flips are cued rather than the belief about unpredictability, or equality of probability.

The conflict involved in choosing between these two basic beliefs is evident in the protocol of S_7, who was asked to justify her initial answer that the coin was more likely to come up tails in the four-heads problem than another heads. She began reasoning from the premise that samples of coin flips will not contain long runs—in other words, that a sample will not contain excesses of one outcome over another. But this reasoning led her to a conclusion that conflicted with her belief that the probability of heads and tails in coin flipping is 50/50.

> *S_7:* T is more likely than another H, because if it's a fair coin, you can only have so many H's in a row. Like, the probability gets smaller of getting another H the more times you flip it, you know, in a row. So I would say that T is more likely.
>
> *I:* You hesitated?
>
> *S:* I hesitated because the outcome of H and T is equally likely, that could be true, too. I'm not sure, because if you look at each separate—that's a separate thing. Me flipping the coin has nothing to do with how I flipped it before. So it might be *c* [equally likely], I'm not sure.

I: So what do you think it probably is?

S: Now that I think about it, probably *c*, because the last—the fifth flip has nothing to do with the first four. It has equal—it's *c*, I would imagine.

She apparently resolves the conflict by reference to a third belief, that each trial is independent of ("has nothing to do with") previous flips.

A related conflict is apparent in the protocol of S_{10}, who finds a different resolution.

S_{10}: Well, for that particular flip the outcomes are equally alike. But if you're going to talk out of all five, then tails is more likely to come up. You know what I mean? So I don't know if you mean that particular flip or....

I: We're talking about that flip.

S: OK, it would just—the outcomes are equal.

I: After it's been flipped four times?

S: They're equally likely for that one.

I: OK. And why is that?

S: It's a 50/50 chance for the coin.

I: OK.

S: But if you're talking about, like one out of five times and, you know, four out of five were heads, the probability would be tails for the fifth one, but I don't—you know what I mean?

On the basis of the last response, this subject was categorized as believing that tails was more likely on the fifth flip. When later asked the probability of getting heads, this subject responded:

S_{10}: Now, see, I still don't understand. Are you, like, saying that....

I: I'm talking about that specific time, when you flip it.

S: It would just be .5.

I: And then for a tails it would be?

S: .5.

I: OK.

S: For each individual one, definitely .5.

I: OK.

S: But ... if you compare it to, you know. If you get four heads in a row, and you say, "What's the probability of getting a tails on the next one?" it would be higher, I think. Like—you know what I mean? Like, just if it's .5 to go either way, then most likely it's going to be the tails on the next flip. That's what I thought you meant at first.

I: OK. I'm having problems understanding how that question is different from this question.

S: Well, if you just say "I'm going to flip the coin," and you tell me the probability of getting heads, I'd say .5. But if you said, "What's the probability that I get heads for the fifth time in a row?" that's, that's what I'm saying.

Our interpretation is that this subject wants to know whether the problem is asking about a single flip of the coin, in which case the probability of coming up heads is 1/2, or whether it is asking about a sample of four flips, in which the first five flips are known to have come up heads. In the latter case, he asserts that "most likely it's going to be the tails." The distinction he makes suggests a conflict between a belief that the coin has an equal chance to come up either heads or tails and the belief that in a sample of flips one expects roughly half heads and half tails. His resolution is to distinguish between what happens in single trials, where he can maintain the equal-probability proposition, and what happens in samples, where the number of heads and tails must be equalized. Asking him how the coin would know the results of the previous four flips might have put him back into a state of conflict.

Although we have presented the ideas of reasoning from general frameworks and reasoning from maxims as two separate accounts, they are not incompatible. Reasoning about uncertain events can be thought of as guided at various times by microstructures, such as the beliefs just cited, and at other times by macrostructures, such as the representativeness heuristic or the outcome approach. In either case, we are suggesting that there is a pattern to the inconsistencies among a student's answers about some chance event, that they are neither a reflection of basic deficits in logical reasoning nor a result of simple carelessness.

CONCLUSIONS

A few important instructional implications follow from the view that people reason from a variety of frameworks and maximlike beliefs about uncertainty. First, assessments of probabilistic or statistical reasoning that are based on correct performance on a few multiple-choice items are not necessarily indicative of a normative understanding. For example, it would be easy to conclude from subjects' performance on the Four-heads problem and Part 1 of the HT-sequence problem that the majority of subjects believe that successive trials in coin flipping are independent and that, therefore, all possible sequences are equally likely. This is the conclusion that would seem to follow from the results of the fourth administration of the National Assessment of Educational Progress (NAEP). Asked for the most likely outcome of a fair coin flip, given four successive trials on which the coin landed tails up, 47% of the 7th graders and 56% of the 11th graders selected the correct alternative. The percentage of responses that were incorrect but consistent with the representativeness heuristic was 38% for the 7th graders and 33% for the 11th graders (Brown et al., 1988). Given that probability is infrequently taught at the secondary school level, these data suggest that most students have a well-developed concept of independence prior to any formal instruction. This conclusion is further supported by results reported by Green (1982). Our research, however, suggests that a sizable percentage of correct responses to such problems are spurious and reflect an outcome approach to uncertainty that is perhaps more pernicious than misapplication of the representativeness heuristic. Problems need to be developed that can discriminate individuals who reason according to the outcome approach from those with a more developed concept of independence.

The belief that the majority of novices faced with these types of problems will commit the gambler's fallacy has helped to shift the focus in probability instruction away from computational skills toward conceptual development (cf. Garfield & Ahlgren, 1988). This shift has been accompanied by curricula aimed at the development of concepts such as independence and randomness and items designed to test for conceptual understanding. Curricula designed by Shaughnessy (1977), Beyth-Marom and Dekel (1983), and Konold (1991a), for example, include units intended to confront and correct judgments based on informal judgment heuristics. A prototypical lesson involves having students make predictions about coin flipping and then test those predictions by drawing several large samples in a computer simulation. The hope has been that those who make predictions on the basis of the representativeness heuristic will be convinced of the fallibility of their intuition when their predictions are shown to differ from the results of the simulation.

Two observations follow from our study about this basic approach to curriculum design. First, if reasoning about various chance events is often based on maximlike beliefs rather than on more general frameworks, such as representativeness or the outcome approach, it might prove fruitful to design early experiences pitting expectations based on several of these maxims against one another. These experiences might promote the need for an integrated, coherent theory of probability.

Second, the curricula mentioned above were designed with the assumption that one way to produce conceptual change is to create situations for which the answers based on a particular incorrect intuition produce cognitive conflict. The resolution of this conflict ideally requires the student to formulate a more normative understanding of the situation (cf. Lakatos, 1976; McDermott, 1984; Minstrell, 1982). The results of the present study suggest one limitation to the cognitive-conflict approach—a situation designed to contrast normative with informal reasoning may produce no conflict. From a normative perspective, the answers some students give to Parts 1 and 2 of the HT-sequence problem are contradictory and thus might be expected to generate the kind of conflict that could promote conceptual change. However, there is little indication that subjects who gave these answers experienced conflict. The important point is that what may appear to the instructor to be a contradiction may nevertheless induce little or no conflict in the student. This does not mean, as is sometimes assumed, that the student has a basic deficiency in logical reasoning. Instead, the lack of conflict may occur because, from the student's perspective, there is no contradiction to be perceived. Incompatibilities and contradictions at this level will probably not be noticed by students unless they are reasoning from a single, coherent framework.

The picture that is emerging from research on student conceptions of probability is that there is no simple story about how students reason about chance. Indeed, one of the major reasons that probability is notoriously difficult to teach is that students bring into the classroom not just one but a variety of beliefs and perspectives about chance. Believing, as we do, that student conceptions must be addressed in the process of instruction, we also acknowledge that a formalized curriculum cannot address all of these. Thus, it is important for teachers of probability to become familiar with the variety of alternative conceptions. In our experience (e.g., Konold 1991b), teachers become more effective as they increase their power to interpret student utterances, many of which may initially seem incomprehensible. Therefore, teachers of probability would do better to minimize, or perhaps delay, their role as a producer of cognitive conflict, adopting instead the ethnographer's frame, trying to understand the language and practices of a foreign culture.

REFERENCES

Beyth-Maron, R., & Dekel, S. (1983). A curriculum to improve thinking under uncertainty. *Instructional Science, 12,* 67–82.

Brown, C. A., Carpenter, T. P., Kouba, V. L., Lindquist, M. M., Silver, E. A., & Swafford, J. O. (1988). Secondary school results for the fourth NAEP mathematics assessment: Discrete mathematics, data organization and interpretation, measurement, number and operations. *Mathematics Teacher, 81,* 241–248.

Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology, 18,* 253–292.

Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education, 19,* 44–63.

Green, D. R. (1982). *Probability concepts in 11–16-year-old pupils.* Centre for Advancement of Mathematical Education in Technology, Loughborough University, Loughborough, United Kingdom.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3,* 430–454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251.

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

Konold, C. (1991a). *Chance: An introduction to probability through simulations.* Unpublished manuscript, Scientific Reasoning Research Institute, University of Massachusetts, Amherst.

Konold, C. (1991b). Understanding students' beliefs about probability. In E. von Glasersfeld (Ed.), *Radical constructivism in mathematics education* (pp. 139–156). Dordrecht, Netherlands: Kluwer.

Lakatos, I. (1976). *Proofs and refutations.* Cambridge: Cambridge University 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. *Physics Teacher, 20,* 10–14.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review, 90,* 339–363.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice-Hall.

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

Shaughnessy, J. M. (1977). Misconceptions of probability: An experiment with a small-group, activity-based, model-building approach to introductory probability at the college level. *Educational Studies in Mathematics, 8,* 295–316.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465–494). New York: Macmillan.

Slovic, P., Fischoff, B., & Lichtenstein, S. (1976). Cognitive processes and societal risk taking. In J. S. Carroll & J. W. Payne, (Eds.), *Cognition and social behavior.* Hillsdale, NJ: Lawrence Erlbaum.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Review, 76,* 105–110.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211,* 453–458.

AUTHORS

CLIFFORD KONOLD, Director of Research, Scientific Reasoning Research Institute, University of Massachusetts, Amherst, MA 01003

ALEXANDER POLLATSEK, Professor, Department of Psychology, University of Massachusetts, Amherst, MA 01003

ARNOLD WELL, Professor, Department of Psychology, University of Massachusetts, Amherst, MA 01003

JILL LOHMEIER, Graduate Research Assistant, Department of Psychology, University of Massachusetts, Amherst, MA 01003

ABIGAIL LIPSON, Senior Psychologist, Bureau of Study Counsel, Harvard University, 5 Linden Street, Cambridge, MA 02138