The Origin of Inconsistencies in Probabilistic Reasoning of Novices

Clifford Konold, Alexander Pollatsek, Arnold Well and Jill Hendrickson - Amherst, Massachusetts, USA
Abigail Lipson - Cambridge, Massachusetts, USA

1. Introduction and method

This paper describes a study in which subjects were asked about various aspects of coin flipping. Many gave contradictory answers to closely-related questions. We offer two explanations for such responses: (a) switching among incompatible perspectives of uncertainty, including the outcome approach (Konold, 1989), judgement heuristics (Tversky and Kahneman, 1982), and normative theory; and (b) reasoning via basic beliefs about coin flipping. As an example of the latter, people believe both that a coin is unpredictable and also that certain outcomes of coin flipping are more likely than others. Logically, these beliefs are not contradictory; they are, however, incomplete. Thus, contradictory statements (and statements at variance with probability theory) appear when these beliefs are applied beyond their appropriate domain.

In the study, twenty subjects (twelve women and eight men) were recruited from undergraduate psychology courses at the University of Massachusetts. Eleven of the subjects had taken, or were currently enrolled in, a statistics course. Subjects participated in an hour-long videotaped interview (conducted by Hendrickson) that included several questions concerning various aspects of probability. The questions about coin flipping came in the middle of the interview. These questions are presented below in the order in which they were asked.

4-Heads Problem: "A fair coin is flipped 4 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?"

H/T Sequence Problem: 1. "Which of the following sequences is *most* likely to result from flipping a fair coin 5 times? (a) HHHTT; (b) THHTH; (c) THHTT; (d) HTHTH; (e) all four sequences are equally likely." 2. "Which of the following sequences is *least* likely to result from flipping a fair coin 5 times?" [The options above were repeated, with option (e) worded "... equally unlikely."]

Probes: After giving answers to these questions, subjects were asked to explain their answers and to give probability estimates for various sequences. Almost all the interviewer's probes (which are available upon request) were standardised in a series of pilot interviews. However, the interviewer was free to ask questions that were not planned when, in her judgement, further clarification was required.

TABLE 1

Correct and incorrect responses to the 4-Heads and H/T Sequence problems

Group	Subject	4-Heads Problem: Equal	4-Heads Problem: P=.5	H/T Sequence Problem: Most	H/T Sequence Problem: Least	H/T Sequence Problem: P =	H/T Sequence Problem: Psum
	6*	+	+	+	+	+	+
	16	+	+	+	+	+	+
A	20	+	+	+	+	+	+
	13	+		+	+	+	+
	17	+	+	+	+	?	?
B	2*	+		+	+		
	15	+	+	?	?	?	?
	12*	+	+	+	+	+	-
	14*	+	+	+	-	+	+
	19	+	+	+	+	-	+
	8*	+		+	+	-	+
	4*	+	+	-	-	+	+
	5	+	+	+	-	+	-
C	11*	+	+	+	-	-	+
	18	+	+	+	+	-	-
	9*	+		-	-	+	+
	3	-	+	+	-	-	-
	7*	+	+	-	-	-	-
	1*	-	+	-	-	-	-
D	10*	-	-	-	-	-	+
% Correct		85	94	74	53	53	65

Table 1 codes correct (+) and incorrect (-) responses for each subject on six aspects of the interview questions; (?) means that the subject did not think it was possible to answer, and a blank indicates a missing value. Subject numbers are listed down the left of the table, response categories along the top. The subjects are in groups A, B, C, D. A (*) indicates prior or concurrent attendance at a probability/statistics course.

The first two columns indicate subjects' answers to the 4-Heads problem. The first column (Equal) indicates whether they correctly answered that the two outcomes were equally likely, and the second (P = .5) whether they gave a probability value of heads (or tails) as .50. The last four columns note correct and incorrect responses to questions that were asked as part of the H/T Sequence problem. Columns three (Most) and four (Least) indicate correct and incorrect responses to the most- and least-likely versions of the problem. A plus in column five (P =) indicates that the probabilities a subject gave for the various sequences were equal. A plus in column six (Psum) indicates that the sum of probabilities given for the mutually exclusive sequences was less than 1. The percentage 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.

2. Salient inconsistencies

Of primary interest in this study was the consistency of a subject's responses over problems. To convey an overall impression of this aspect, subjects are ordered in Table 1 according to total number of correct responses. Subjects listed in the top group, A, of the table answered all of the problems normatively. The subject in the bottom group, D, answered consistently in accord with the representativeness heuristic (Tversky and Kahneman, 1982). Subjects in group B showed no obvious inconsistencies, but did not respond to a number of questions. The 12 subjects in group C showed various inconsistencies in their responses and thus are of greatest interest to us here. The most salient of these inconsistencies are described below.

(i) *H/T Sequence : most- vs least-likely outcomes.* Subjects who respond that all sequences in the H/T Sequence problem are equally likely should then respond, if they are reasoning normatively, that all sequences are equally *un*likely. However, subjects S_3, S_5, S_{11}, and S_{14} each responded that all the sequences in the H/T Sequence problem were equally likely but went on to choose a sequence as *least* likely.

(ii) *Qualitative answers vs probabilities : H/T Sequence.* Subjects who respond that all of the sequences are equally likely ought then to give equal probabilities to the options. However S_8, S_{18}, and S_{19} each responded that all of the sequences were equally likely and then assigned them unequal probabilities.

Similarly, subjects who choose a particular sequence as most (or least) likely ought to assign a greater (or lesser) probability to that sequence than to the other sequences. However, 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).

(iii) *Qualitative answers vs probabilities : 4-Heads.* Subjects reasoning normatively about the 4-Heads problem will respond that heads and tails are equally likely and then will assign equal probabilities to each outcome. Although most of the subjects responded accordingly, S_1 and S_3 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.

(iv) *Responses on the 4-Heads vs H/T Sequence problems.* Subjects correctly answering the 4-Heads problem would seem to be exhibiting an understanding of independence of successive trials in coin flipping. Given such an understanding, these subjects ought to regard the various sequences in the H/T Sequence problem as equally likely. S_4, S_7 and S_9 all gave correct answers to the 4-Heads problem but incorrect responses to both versions of the H/T Sequence problem.

(v) *Constraint on the sum.* Six subjects gave probability values for a small subset of sequences whose sum equalled or exceeded 100%. It should be noted that since subjects were asked for the probabilities of only 2 or 3 of the total 32 sequences, these results provide a conservative estimate of subjects' misunderstanding of the constraint on the sum of probabilities of mutually exclusive events. S_4, for example, gave 10-20% as the probability range for each of the options "b", "d", and "c" and is coded in Table 1 as responding correctly. Had she been asked for the probabilities of a few more sequences, she undoubtedly would have continued giving the same probability range and thus wold have exceeded the sum of 100%.

3. Possible explanations for inconsistent responses

On a few occasions, subjects noticed that their responses seemed contradictory, but their attempts to describe the nature of that confusion were not particularly coherent. In a few other cases, subjects may have been aware of inconsistencies but were reluctant to verbalise their confusion. For the most part, subjects indicated no awareness of inconsistencies in their responses. In this section we propose two general explanations of the inconsistencies mentioned above.

(i) *Switching among alternative perspectives of uncertainty.* One possible explanation for inconsistencies is that subjects hold multiple frameworks about probability, and subtle differences in situations activate different perspectives. Moreover, different perspectives can be employed almost simultaneously in the same situation. Framework switching may explain, for example, inconsistencies between subjects' responses to the most- and least-likely versions of the H/T Sequence problem. People reasoning according to the outcome approach interpret a question about the probability of an event as a request to predict whether or not that event *will* occur. These people may choose answer "e" ("equally likely") to indicate that *a prediction cannot be made*. For example, S_5 gave the following justification for her response of "equally likely":

> S_5: "You can't tell. It's a game of chance. 50/50 ... it could be anything."

But in the least-likely version one cannot interpret the question as a request to predict which one sequence will not occur. This is because all (or all but one) of the sequences will not occur. A paraphrasing of subjects' thinking about these two different versions of the problem might be "I can't say which sequence will occur [most-likely version], but I think sequence x is particularly unlikely [least-likely version]." S_3 came closest to this justification:

> S_3: "[Most likely: e] I just think that they - any of them could happen, cause it's not like all heads or all tails.
> [Least likely: c] Because tails would have to come up the last 3 times."

Framework-switching may also account for the inconsistencies that several subjects demonstrated between their qualitative answers and their assignment of probabilities. In the H/T Sequence problem, for example, the subjects described above may respond in accord with the outcome approach that all the sequences are equally likely and regard it as consistent to assign higher probabilities to some of the outcomes than to others by virtue of the representativeness heuristic. Again, the task "which sequence is most likely" may be interpreted as a request for a prediction, while the task of assigning probabilities is not. Because the probabilities are not different enough, these subjects may not regard the differences as providing a sound basis for prediction.

(ii) *Conflict among basic beliefs about coin flipping.* Inconsistencies in subjects' reasoning may also be accounted for in terms of conflicts among basic beliefs associated with flipping coins. In this account, inconsistencies result from conflicts among these beliefs rather than from switching among different probability frameworks. *Some* of the assertions that subjects made repeatedly during the interview included the following:

(a) One cannot predict for certain the results of coin flipping.
(b) The outcome of repeated trials varies erratically between heads and tails.
(c) A coin has no memory.
(d) Heads and tails are equally likely.
(e) Heads and tails occur about equally often in a sample of flips.

Note that all of these beliefs are correct. They are basic in the sense that they are used to justify other statements and are, themselves, self-evident.

The inconsistent responses by S_4, S_7, and S_9 to the 4-Heads and H/T Sequence problems may be an example of reasoning from these basic beliefs. The 4-Heads problem asks subjects about the probability of heads vs tails conditioned on the results of four previous flips. The primary beliefs that this problem seem to cue are that coin flipping is unpredictable and that the coin has no memory. In fact, the term "50/50" seems to be used as a synonym for "unpredictability". The fact that the coin is about to be flipped once may be more salient to subjects than the two resultant sequences HHHHT and HHHHH. Accordingly, the belief in unpredictability is cued rather than the belief that samples ought to contain roughly equal numbers of heads and tails. In the H/T Sequence problem, on the other hand, the sequences are more salient than single flips. As a result, beliefs in the equality of heads and tails in samples and the irregularity of flips are cued rather than the belief about unpredictability.

While we have presented the ideas of reasoning from general frameworks and reasoning from basic beliefs as two separate accounts, they are not incompatible. Reasoning about uncertain events can be thought of as guided at various times by micro structures, such as the beliefs just cited, and at other times by macro structures, such as representativeness and the outcome approach.

4. Conclusion

Two important teaching implications follow from the view that people use a variety of frameworks and beliefs concerning 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 4-Heads problem and the most-likely version of the H/T Sequence problem that most subjects believe that successive trials in coin flipping are independent, and that therefore all possible sequences are equally likely. Similar problems were used by NAEP in assessing probabilistic reasoning of a national sample of secondary school students in the United States (Brown et al., 1988). The high rate of correct responses on these items could wrongly convince educators that the gambler's fallacy was well on its way to becoming obsolete.

Secondly, an instructional technique that is gaining popularity in domains in which students hold strong misconceptions, such as physics, is to have students collect data that are inconsistent with their intuitive predictions. The idea is to motivate students to abandon theories that do not stand up to empirical scrutiny, substituting these eventually with theories that are more successful predictors. A similar approach has been suggested in teaching probability and statistics and has become especially attractive with the availability of computers in the classroom. A prototypical lesson involves having students make predictions about coin flipping and then testing those by drawing several large samples in a computer simulation. The fantasy is that those who make predictions on the basis of the representativeness heuristic are quickly convinced about the fallibility of their intuition.

However, if people reason about coins from a variety of perspectives, and can make predictions that are incompatible, such simulations may have little impact on beliefs. Students may as easily switch among different perspectives in trying to account for a particular set of data as they do in making predictions of what will happen. The incompatibility among both their predictions and their observations will not be noticed by students until they are reasoning from a single framework.

Acknowledgement and Disclaimer

This reseach was supported by Grant MDR-8954626 from the National Science Foundation. The views expressed here to not necessarily reflect those of the Foundation.

References

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

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

Tversky, A and Kahneman, D (1982) Judgment under uncertainty : heuristics and biases. In: D Kahneman , P Slovic and A Tversky (eds) *Judgment Under Uncertainty : Heuristics and Biases.* Cambridge University Press, New York, 3-20.