Partition–Edit–Count: Naive Extensional Reasoning in Judgment of Conditional Probability

Craig R. Fox
University of California, Los Angeles

Jonathan Levav
Columbia University

The authors provide evidence that people typically evaluate conditional probabilities by subjectively *partitioning* the sample space into *n* interchangeable events, *editing* out events that can be eliminated on the basis of conditioning information, *counting* remaining events, then reporting probabilities as a ratio of the number of focal to total events. Participants' responses to conditional probability problems were influenced by irrelevant information (Study 1), small variations in problem wording (Study 2), and grouping of events (Study 3), as predicted by the partition–edit–count model. Informal protocol analysis also supports the authors' interpretation. A 4th study extends this account from situations where events are treated as interchangeable (chance and ignorance) to situations where participants have information they can use to distinguish among events (uncertainty).

People are often called on to make judgments of conditional likelihood. For example, a patient might try to assess the likelihood of having a disease, given a positive test result; a litigant might try to estimate the odds of prevailing in court, given a piece of damning evidence. Over the last 30 years, psychologists have shown that people typically judge conditional probabilities using a limited set of heuristics that reduce these tasks to more basic assessments of similarity or memory accessibility (Gilovich, Griffin, & Kahneman, 2002; Kahneman, Slovic, & Tversky, 1982). For instance, one might judge the likelihood that a person is a bank teller, given a description of that person, by assessing the similarity between the description and one's impression of the prototypical bank teller; one might judge the proportion of crimes committed by women by assessing the ease with which examples of crimes by women come to mind relative to the ease with which crimes by either women or men come to mind (Tversky & Kahneman, 1983).

Heuristics such as representativeness and availability entail an evaluative assessment of the nature of target events. The purpose of this article is to characterize and test a second, complementary strategy for judging conditional probability that entails an assessment of the number of events that could occur. This strategy is based on the definition of probability as a ratio of interchangeable events.[1] Although the formal theory of probability is a human invention that did not coalesce until the 17th century, people apparently acquire an intuitive conception of probability without formal schooling on the topic. Piaget and Inhelder (1951/1975) reported that by age 10 or 11 children come to understand chance probabilities as the proportion of favorable cases to total (favorable plus unfavorable) cases. More recent work suggests that children may acquire this intuitive notion as early as 9 and develop a capacity to explicitly match probability ratios by age 13 (Falk & Wilkening, 1998; see also Girotto & Gonzalez, 2003).

We argue that judgments of conditional probability entail a combination of evaluative assessment and naive extensional reasoning concerning the number of events that could occur. Reliance on this latter strategy is most apparent in situations where there is a set of possible events that cannot be distinguished readily through the use of evaluative strategies. Consider the following example:

> Your department decides to interview three qualified candidates (Ames, Boyd, and Clark), in an arbitrary order, for a junior faculty opening. After Ames and Clark have visited, the faculty meet and agree on the following: (1) before interviewing any candidates there was no consensus concerning who was strongest; (2) after seeing Ames and Clark, Ames was judged to be stronger than Clark; (3) if, on the basis of (1) and (2), there is at least a 60 percent chance that Ames is the *strongest* of the three candidates, the department will make an immediate offer to Ames. *What is your best estimate of the probability that Ames is the strongest of the three candidates, given that Ames is stronger than Clark?*
>
> (Please take a moment to decide before reading further.)

---

[1] Our term *interchangeable events* should be distinguished from de Finetti's (1931, 1937) term *exchangeable events*, which refer to "events that occur in a sequence in a random order; the order of their occurrence does not affect the probabilities we are interested in" (Kyberg & Smokler, 1980, p. 17).

In this case, intensional considerations of similarity or availability of instances do not seem relevant. Instead, we assert that the identification of interchangeable events (three candidates interviewed in a random order) leads most people to invoke a naive form of extensional reasoning: (a) initially each candidate had an equal chance of being the strongest; (b) learning that Ames is a stronger candidate than Clark allows us to rule out Clark as the strongest candidate; (c) hence, two equally likely possibilities remain—either Ames or Boyd is the strongest candidate, so that the probability that Ames is strongest is 1/2, and therefore this candidate should not be offered the job.

Although such reasoning may appear compelling, it is incorrect in this case because learning that Ames is stronger than Clark also provides some information concerning the relative ordering of Ames and Boyd.[2] Let A denote the event "Ames is the strongest candidate," and let $A > C$ denote the information "Ames is a stronger candidate than Clark." Note that at the outset these probabilities are $P(A) = 1/3$ and $P(A > C) = 1/2$; because the latter event is true whenever the former event is true, we also know that $P[A \& A > C)] = 1/3$. According to the definition of conditional probability,

$$P(A|A > C) = \frac{P[A\&(A > C)]}{P(A > C)} = \frac{1/3}{1/2} = 2/3. \qquad (1)$$

Hence, the correct answer is 2/3, and the department should indeed make an immediate offer to Ames.

Note that one can also derive the correct answer to the Ames–Boyd–Clark problem using the more intuitive notion of probability as a ratio of numbers of events—provided the events under consideration remain equiprobable after one learns the conditioning information. If one partitions the set of possibilities not by which candidate is strongest {A, B, C} but instead by the possible rankings of candidates, one gets {ABC; ACB; BAC; BCA; CAB; CBA}, with each ordering equally likely at the outset (ABC is interpreted as $A > B > C$, and so forth). Learning that $A > C$ allows us to eliminate the last three orderings, in which $C > A$. This leaves three orderings that remain equiprobable,[3] two of which rank A ahead of B, yielding the correct answer of 2/3.

In this article, we explore the role of naive extensional reasoning in conditional probability assessment. We focus our attention primarily on situations in which elementary events are treated as indistinguishable, so that evaluative approaches such as representativeness and availability do not apply. The study of chance illusions can help us understand how people process probabilistic information, just as the study of visual illusions has shed light on how people process visual information.

In our experiments we adapt well-known probability puzzles that share a common structure: (a) Basic parameters are described (e.g., a set of job interviews is scheduled in an arbitrary order), (b) new information is provided (e.g., Ames is stronger than Clark), and (c) a conditional likelihood judgment is elicited (e.g., what is the probability that Ames is the strongest candidate, given that Ames is stronger than Clark?). We conjecture that people typically solve these problems by appealing to the most basic and intuitive notion of probability as a ratio of interchangeable events. In particular, we propose that people (a) subjectively *partition* the sample space into a set of *n* interchangeable events suggested by the parameters provided (e.g.,{Ames is strongest; Boyd is stron-

gest; Clark is strongest}), (b) *edit* out events that can be eliminated on the basis of the new (conditioning) information (e.g., learning that Ames is stronger than Clark allows us to eliminate "Clark is strongest" from consideration), and (c) *count* the number of target events and total events that remain, and then report the relevant ratio or quotient (e.g., "Ames is strongest" is one of two remaining possibilities, so the probability is 1/2).

We propose further that people normally invoke the simplest partition of the sample space that can accommodate the basic parameters but that they are more likely to invoke a more refined partition when it is made more accessible through variations in the description of the problem (e.g., asking about rankings of all candidates rather than who is strongest). Therefore, we argue that judgments of conditional probability are *partition dependent,* varying as a function of the (edited) partition of the sample space that a person invokes. The notion that different descriptions of the same problem facilitate different psychological representations—and, in turn, different responses—has been documented previously in the domains of logical reasoning (Wason & Johnson-Laird, 1972), problem solving (Kotovsky & Simon, 1990; Kotovsky, Hayes, & Simon, 1985), decision making (Tversky & Kahneman, 1986), and judgment under uncertainty (Fox & Rottenstreich, 2003; Rottenstreich & Tversky, 1997). We now extend this tradition to the domain of naive extensional reasoning.

In the studies that follow, we provide evidence that people rely on the *partition–edit–count* strategy in conditional probability assessment. Our basic method is to vary the nature, modality, or description of information presented in ways that do not affect normative probability but are designed to influence the relative accessibility of partitions or edits. Studies 1A and 1B explore editing using a simple card game adapted from the so-called Monty Hall problem. Studies 2A, 2B, and 2C explore partitioning using variations of classic probability puzzles. Study 3 demonstrates that a person's partition of the sample space is not necessarily determined by the physical wholeness of objects being sampled (cf. Brase, Cosmides, & Tooby, 1998) but can be influenced by whatever ad hoc grouping happens to be made accessible to the judge. Study 4 provides evidence that the partition–edit–count strategy may be used in combination with evaluative strategies in situations in which people are able to distinguish the nature of elementary events. We conclude with a general discussion of our results, including an account of related work, a formal approach to modeling the partition–edit–count strategy, and implications for the instruction of elementary probability theory.

---

[2] To convince oneself, one could imagine that the department had interviewed 10 candidates and that Ames was the strongest of the first 9. Clearly, Ames must be a very strong candidate, so his chances of being stronger than the 10th candidate are better than 50–50.

[3] If this problem is to be solved correctly, the partition must be sufficiently refined so that it can incorporate all relevant information provided by the conditioning event. Learning that $A > C$ tells us not only that C is not the best candidate but also that if B is best, then $A > C$. Hence, B is best in only one possible ordering (BAC), but A is best in two possible orderings (ABC or ACB). A partition by strongest candidate, {A, B, C}, therefore, does not allow us to represent probability as the ratio of cases remaining, whereas a partition by order {ABC, ACB, BAC, BCA, CAB, CBA} does.

Study 1: Investigations of Editing

Perhaps the most notorious probability puzzle that has been investigated in recent years is the Monty Hall problem. It came to national attention when *Parade* magazine columnist Marilyn vos Savant (1990) published the following question in her weekly column, based on Monty Hall's television show "Let's Make A Deal":

> Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

Gardner (1961) had reported an isomorphic problem concerning three prisoners in a column that appeared many years earlier in *Scientific American*. The problem has fascinated statisticians, mathematicians, psychologists, and laypeople for years because its solution is wildly counterintuitive (see, e.g., Ben-Zeev, Stibel, Dennis, & Sloman, 2003; Falk, 1992; Franco-Watkins, Derks & Dougherty, 2003; Friedman, 1998; Granberg & Brown, 1995; Kraus & Wang, 2003). Most people, when presented with this problem, are convinced that the probability of the prize being behind Door Number 2 is 1/2, whereas, in fact, the probability is 2/3. In our first study we modified the original problem so that we could eliminate some informational ambiguities inherent in its original wording (cf. Morgan, Chaganty, Dahiya, & Doviak, 1991; Tor & Bazerman, 2003). We developed a game of chance that allowed us to manipulate the accessibility of various potential edits to the sample space without changing the normative probability. Our hypothesis is that such modifications will influence participants' assessments of probability and their associated choices.

Study 1A: Five Card Monty: Influencing the Editing Operation

Method

We recruited 104 visitors to the Duke University student center through posters in the building that offered participants a chance to win $1 for completing a 5-min task. Each participant was shown a deck of five playing cards that included an ace, deuce, three, four, and five. The experimenter shuffled the cards until the participant indicated that he or she was satisfied that the cards had been randomized properly. Next, the experimenter dealt two cards to the participant and three cards to himself, all face down. The participant was told that if one of his or her cards at the end of the game was the ace, he or she would be awarded $1. The present account suggests that the problem should initially facilitate a fivefold partition of the sample space, with each card assigned an equal chance of being the ace. This setup was followed by an experimental manipulation that was designed to facilitate various edits of the fivefold partition. Participants were assigned at random to one of the following three experimental conditions:

1. No cards up (*n* = 36): "I'm going to look at my hand, then place all three cards back down. After that I will ask you if you want to trade your cards for my cards."

2. One card up (*n* = 35): "I'm going to look at my hand, then turn up one card that is not the ace. After that I will ask you if you want to trade your cards for my cards."

3. Two cards up (*n* = 33): "I'm going to look at my hand, then turn up two cards that are not the ace. After that I will ask you if you want to trade your cards for my cards."

Next, the experimenter looked at the cards in his hand and then turned up the number of cards promised (where relevant, we counterbalanced the particular nonace cards that were turned up). The participant was then asked to state his or her best estimate of the probability that the ace was (one of) the experimenter's face-down card(s) rather than one of the participant's two cards. Finally, the participant was asked whether he or she wanted to trade his or her cards for the experimenter's cards.[4] After this choice was made, all cards were revealed, and the participant received $1 if the ace was in his or her hand.

The probability that the ace is initially dealt to the experimenter is 3/5 = .60, and of course this probability does not change in light of the new information, because the experimenter can always find at least two cards in his hand that are not the ace. Hence, from a normative standpoint, all participants should choose to switch cards with the experimenter. The present account predicts that participants will answer the problem correctly when no cards are turned face up, because this is a simple problem in which there are five interchangeable events and probability can be calculated as a simple ratio of cases. However, when the experimenter turns *n* nonace cards in his hand so that they are facing up, many participants will edit these cards from their subjective partition so that the judged probability of the target event is $(3 - n)/(5 - n)$, and the proportion of participants switching cards will decrease as *n* increases.

Results and Discussion

The results closely follow the predicted pattern of editing, consistent with use of the partition–edit–count strategy (see Table 1). Median and modal judged probabilities were 3/5, 1/2 (i.e., 2/4), and 1/3 when the experimenter turned up no cards, one card, and two cards, respectively. In fact, we received these responses from a strong majority of participants in these three conditions, and the proportions of responses of 3/5, 1/2, and 1/3 varied significantly by experimental condition, $\chi^2(4, N = 71) = 112$, $p < .0001$.[5] The proportion of participants willing to trade also decreased monotonically with the number of cards turned up, as predicted. The overall difference in the proportion of participants willing to trade across experimental conditions was highly significant (see Table 1, Column 5), $\chi^2(2, N = 104) = 16.91$, $p < .001$. Despite the fact that participants in all experimental conditions should have traded their cards for the experimenter's card(s) to increase their probability of winning from 2/5 to 3/5, very few participants chose to do so. Even in the no cards up condition, in which a large majority of participants correctly judged the probability of the prize being in the experimenter's hand to be 3/5, a bare majority opted to trade cards. This overall reluctance to trade might be due to asymmetric regret, which has been observed in previous studies of the Monty Hall problem: People anticipate that they would experience greater regret if they missed the prize because of their own action (trading

---

[4] Approximately half the participants were asked to make a choice before judging the probability. This order manipulation had no effect on the results, so the data were pooled across ordering conditions.

[5] Franco-Watkins et al. (2003, Study 3) independently devised a similar problem in which the number of both doors and prizes varied across experimental conditions. Judged probabilities in their variation also accord with predictions of the partition–edit–count model.

Table 1
Results of Study 1A

Experimenter's hand	n	Median judged probability	% reporting median	% trading
No cards up	36	.60	81	58
One card up	35	.50	60	23
Two cards up	33	.33	64	15

Note. % trading indicates the percentage of participants who opted to trade their hand for the experimenter's hand.

hands) rather than their inaction (retaining the default hand; see Gilovich, Medvec, & Chen, 1995).

Study 1B: Five Card Monty: Manipulating the Accessibility of an Edited Partition

We have interpreted Study 1A as supporting the notion that participants edit cards from their fivefold partition after these cards have been turned up. Apparently, participants found it natural to adjust their assessment of the probability that the ace was in the experimenter's hand after nonace cards were physically revealed. We surmise that the relative accessibility of particular edits is influenced by how concretely the conditioning information is conveyed. When cards are flipped, they are naturally eliminated from consideration because they are physically set apart from the nonflipped cards. In contrast, we expect that participants would find such editing somewhat less natural if conditioning information were presented in a more abstract way that required them to mentally transform their initial (fivefold) partitions.

Method

We recruited a new sample of participants ($N = 126$) at the Duke University student center, as before. Participants were offered a chance to win $1 if they agreed to complete a 5-min task. The procedure was virtually identical to that used in the two cards up condition of Study 1A, with the following adjustments. After dealing two cards to the participant and three to himself, the experimenter told participants, "I'm going to look at my hand, then indicate two cards that are not the ace. After that I will ask you if you want to trade your cards for my cards."

The experimenter then looked at his hand and stated, "These two cards are not the ace." As he said these words, the experimenter either physically turned the indicated cards face up ($n = 64$) or merely pointed to the indicated cards ($n = 62$). Next, the participant was asked to state the probability that the ace was in the experimenter's hand rather than in his or her own hand. Finally, the experimenter offered to switch hands with the participant. Cards were then turned up, and the participant received $1 if the ace was in his or her hand.

Results and Discussion

The results again conformed to our predictions (see Table 2). Replicating the pattern observed in the two cards up condition of Study 1A, most participants judged the probability of the target event to be 1/3 when two nonace cards were physically turned face up; however, fewer than half of the participants committed this error when these cards were not physically turned over (66% vs. 48%), $\chi^2(1, N = 126) = 4.03, p < .05$. As we had predicted, when

the experimenter merely pointed to the cards—thereby keeping the original (unedited) fivefold partition somewhat more salient—participants were more than three times as likely to indicate the correct response of 3/5 (26% vs. 8%), $\chi^2(1, N = 126) = 7.34, p < .01$. Overall, our experimental manipulation was successful in influencing the relative frequencies of responses of 1/3 and 3/5, $\chi^2(1, N = 93) = 7.75, p = .005$. Very few participants opted to switch hands when two cards had been physically turned up, but twice that proportion did so when the experimenter merely pointed to the cards (9% vs. 18%), $\chi^2(1, N = 126) = 1.89, ns$. As with Study 1A, we suspect that the relatively low switching rate in both conditions is a result of asymmetric regret.

Study 2: Investigations of Partitioning

Having provided evidence for the principles of counting and editing, we now focus on subjective partitioning. The Monty Hall problem has several properties that diminish its usefulness for further investigations of partition dependence. First, this problem requires participants to begin by guessing a door (so that the game show host's actions are constrained), which may influence the participant's subsequent probability assessment (e.g., the participant may experience the illusion of control—see, e.g., Langer, 1975). Second, there is an intentional agent (the game show host) with asymmetric information who has discretion over which information is revealed (for discussions of the role of perspective on this problem, see Kraus & Wang, 2003; Tor & Bazerman, 2003). Most important, the problem cannot be solved correctly with the partition–edit–count strategy because there is no obvious partition of the sample space yielding elementary events that are equiprobable in light of the conditioning information (but see Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni, 1999, pp. 82–83; see also Ben-Zeev et al., 2003). The problems that we use in Study 2 suffer from none of these drawbacks.

Our approach is to begin with a version of a well-known probability puzzle and vary the wording slightly without altering its normative solution, in ways designed to make more accessible an alternative partition from which the correct solution can be derived through the use of the partition–edit–count strategy. For each problem, we also asked participants to explain their answers; an analysis of written protocols is presented at the end of this section. To provide a conservative test of the hypothesis that people's answers vary as a function of the superficial wording of each problem, we ran Study 2 using sophisticated participants: master's of business administration (MBA) students who had recently received training in elementary probability theory, includ-

Table 2
Results of Study 1B

Experimenter's hand	n	% responses			% trading
		.33 or 1/3	.60 or 3/5	Other	
Turn cards over	64	66	8	26	9
Point to cards	62	48	26	26	18

Note. The correct response was .60 or 3/5. % trading indicates the percentage of participants who opted to trade their hand for the experimenter's hand.

Table 3
Results of Study 2A (Cancer Drug)

Experimental manipulation				Predicted response		% actual responses by probability			
Target event	Conditioning event	*n*	Primed partition	P-E-C	P-C	1/3	1/2	2/3	Other
A most effective	A more effective than C	67	{A, B, C} naive	1/2	1/3	12	64	10	13
A ranked ahead of B, C	A more effective than C	62	{ABC, ACB, BAC, BCA, CAB, CBA} refined	2/3	2/6 = 1/3	11	53	23	13

Note. P-E-C = partition–edit–count; P-C = partition–count.

ing the computation of conditional probabilities and problems requiring the use of Bayes's theorem. Previous researchers have amply demonstrated errors solving classic problems using more representative cross-sections of participants. We reasoned that more sophisticated participants should, if anything, be less susceptible to errors and less affected by irrelevant changes in wording than participants with only a very basic understanding of the concept of probability.

Study 2A: Cancer Drug

Method

In Study 2A we used a version of the foregoing Ames–Boyd–Clark problem with a cover story that we thought might appeal to our sample of 1st year business students at Duke University (*N* = 129). Students were recruited late in the term from a daytime MBA course in probability and statistics and a weekend MBA class in economics and asked to complete a brief survey.[6] On completion of the survey, some students were selected at random to receive a $20 prize. They were presented the following problem (alternative wordings are listed in brackets):

> Three pharmaceutical companies (A, B, and C) have been developing a new class of cancer-fighting drugs. The FDA has completed a study of the relative effectiveness of all three drugs. On Monday, the FDA will publish a report in which it will [reveal which of the three drugs is most effective/rank the three drugs from most effective to least effective]. You have just learned that an independent laboratory compared the effectiveness of A and C, finding definitively that A is more effective than C. What is the probability that [the FDA will identify A as the most effective of the three?/the FDA's rankings will list A ahead of both B and C?].

Approximately half the respondents (*n* = 67) received the first wording, in which the FDA would "reveal which drug is most effective," then were asked to assess the probability that the FDA would identify Drug A as the "most effective of the three." The remaining respondents (*n* = 62) were told that the FDA would "rank the three drugs from most to least effective," then were asked to assess the probability that Drug A would be "ranked ahead of both B and C."

Of course, our subtle variation in wording does not affect the correct answer to the problem, which is 2/3, just as in the foregoing job candidate anecdote. We conjectured, however, that the "most effective" language would lead participants to invoke a naive, threefold partition of which drug is most effective {A most effective, B most effective, C most effective}, then edit out C in light of the preliminary study, which would yield a probability of 1/2. In contrast, we conjectured that the "ranking" language would make more accessible a refined, sixfold partition of possible rankings: {ABC, ACB, BAC, BCA, CAB, CBA}. In the editing phase, partic-

ipants would then eliminate the latter three possibilities. Because A is ranked ahead of both B and C for two of the three remaining possibilities, this procedure yields a probability of 2/3. Hence, we predicted that respondents would more often arrive at the correct probability of 2/3 when the target query was phrased in terms of rankings rather than which drug would be deemed most effective. Likewise, we predicted that respondents would more often arrive at the incorrect probability of 1/2 when the target query was phrased in terms of which drug would be most effective rather than rankings of the three drugs.

Results and Discussion

The results support our predictions (see Table 3). Respondents were more than twice as likely (23% vs. 10%) to provide the correct response of 2/3 under the "ranking" language than under the "most effective" language, $\chi^2(1, N = 129) = 3.50, p < .06$. In addition, the proportion of responses of 1/2 decreased under the ranking language (from 64% to 53%), though this decrease did not reach statistical significance, $\chi^2(1, N = 129) = 2.07, p = .15$. Overall, our manipulation was successful in affecting the relative frequencies of responses implied by naive versus refined partitions under the partition–edit–count strategy, $\chi^2(1, N = 98) = 3.75, p = .05$.

Further support for the partition–edit–count model can be gleaned from an analysis of responses other than 1/2 or 2/3. Some participants may have invoked the naive or refined partition but failed to condition on the additional information provided (i.e., mistakenly invoked a partition–count strategy without editing; see Column 6 in Table 3). Some of these respondents may have adopted a naive partition in terms of which drug is most effective {A, B, C} but then computed the probability that A was most effective without considering the information that A was more effective than C. This would yield an *unconditional* probability of 1/3. Likewise, some respondents may have adopted the more refined partition and considered the probability that "A is ranked ahead of B and C" without properly conditioning on the information that A was more effective than C. This would also lead to a judged probability of 2/6 = 1/3. Hence, if a sizable minority of participants apply a partition–count strategy without editing, we would expect to see a sizable minority of responses of 1/3 in both conditions. An inspection of Table 3 reveals that nearly half of the

---

[6] Some ancillary details of the cover story were pruned from the original survey so that the second group received the more streamlined version that we present here. There were no significant differences in responses between subsamples, so we combined the data.

participants who did not provide an answer predicted by the partition–edit–count model provided responses of 1/3, and this was the most common alternative response in both conditions.

Study 2B: Mr. Smith's Children

To provide broader evidence for partition dependence, we adapted a second classic problem from Gardner (1961).

Method

We recruited a new sample of MBA students late during the term in a course on probability and statistics at Duke University ($N = 122$) and again selected some respondents at random to receive a $20 prize. Roughly half the participants ($n = 60$) received the following problem: "Mr. Smith says: 'I have two children and at least one of them is a boy.' Given this information, what is the probability that the other child is a boy?"

We submit that this problem is misleading by design because its wording confers privileged status on an arbitrary first child (the boy we learn about), then focuses attention on an arbitrary second child (the child about whose sex we are asked). This facilitates a twofold, naive partition because the "other" child can be either a boy or a girl, and the sex of successive children is independent. Hence, we predicted that these participants would overwhelmingly report a probability of 1/2 (for a related analysis of this problem, see Bar-Hillel & Falk, 1982).

We presented the remaining participants ($n = 62$) a second version of this problem that was designed to convey identical information but facilitate a partition that does not direct attention to an arbitrary child: "Mr. Smith says: 'I have two children and it is *not* the case that they are both girls.' Given this information, what is the probability that both children are boys?"

We expected this new wording to facilitate a fourfold, refined partition that incorporates the sex of both children: {boy–boy, boy–girl, girl–boy, girl–girl}. After the possibility of two girls {girl–girl} is edited from the sample space, it becomes apparent that the correct probability of both children being boys is 1/3. We therefore predicted that the original wording would give rise to a higher proportion of incorrect responses of 1/2 and that the revised wording would give rise to a higher proportion of correct responses of 1/3.

Results and Discussion

Results again conformed to our predictions (see Table 4). Participants were nearly 10 times as likely to indicate the correct response of 1/3 in the revised ("both children") formulation than in the original ("other child") formulation (31% vs. 3.3%), $\chi^2(1, N = 122) = 15.96$, $p < .0001$. Moreover, a strong majority (85%) of participants answered 1/2 under the original problem wording, but only 39% did so under the modified wording, $\chi^2(1, N = 122) = 27.59$, $p < .0001$. Overall, our manipulation was successful in affecting the relative frequencies of responses implied by naive versus refined partitions under the partition–edit–count strategy, $\chi^2(1, N = 96) = 22.69$, $p < .0001$.

Further support for the present account can be gleaned from an analysis of unconditional probabilities that are predicated on naive versus refined partitions. When a participant judges the probability that Mr. Smith's "other child is a boy," the twofold partition yields a probability of 1/2, whether or not the participant considers that "at least one [child] is a boy." However, when a participant judges the probability that "both children are boys," a participant invoking a fourfold partition but failing to incorporate information that "it is not the case that both are girls" will provide a response of 1/4. Consistent with this analysis, responses of 1/4 were much more common under the modified problem wording than the original problem wording (24% vs. 3.3%), $\chi^2(1, N = 129) = 11.06$, $p < .001$.

Study 2C: Three Cards in a Hat

Our final problem is an isomorph of the so-called Bertrand box problem that we adapted from Bar-Hillel and Falk (1982).

Method

We presented the following problem to a sample of executive MBA students ($N = 76$) enrolled in a course in decision models at Duke University; we again selected some respondents at random to receive a $20 prize to reward their participation. Approximately half the respondents ($n = 39$) received the following problem:

> Three two-sided cards are in a hat. One is red on both sides, denoted Red–Red. One is red on one side and white on the other, denoted Red–White. One is white on both sides, denoted White–White. Suppose that you reach into the hat without looking, place a card on the table, open your eyes, and the side showing is *red*. Given that the side showing is red, what is the probability that the hidden side is also red?

When Bar-Hillel and Falk (1982) presented this problem to undergraduate students in a probability course, 66% of their sample answered 1/2, despite the fact that the correct answer is actually 2/3. The authors suggested that participants reasoned that the card was definitely not white–white, so it was either red–red or red–white, which is consistent with the present analysis. That is, people are likely to invoke a naive, threefold partition of the sample space by card {red–red, red–white, white–white}, then edit out the white–white card.

Table 4
Results of Study 2B (Mr. Smith's Children)

Experimental manipulation				Predicted response		% actual responses by probability			
Target event	Conditioning event	n	Primed partition	P-E-C	P-C	1/4	1/3	1/2	Other
Other child is a B	At least one child is a B	60	{B, G} naive	1/2	1/2	3.3	3.3	85	8.3
Both children are Bs	Not the case that both are Gs	62	{BB, BG, GB, GG} refined	1/3	1/4	24	31	39	6.4

Note. P-E-C = partition–edit–count; P-C = partition–count; B = boy; G = girl.

Table 5
Results of Study 2C (Three Cards in a Hat)

Experimental manipulation				Predicted response		% actual responses by probability			
Target event	Conditioning event	*n*	Primed partition	P-E-C	P-C	1/3	1/2	2/3	Other
Other side of card is R	R face up	39	{RR, RW, WW} naive	1/2	1.5/3 = 1/2 or 1/3	13	59	2.6	26
It is side R_1 or R_2	R face up	37	{R_1, R_2, R_3, W_1, W_2, W_3} refined	2/3	2/6 = 1/3	35	24	27	14

Note. P-E-C = partition–edit–count; P-C = partition–count; R = red; W = white.

We presented the remaining respondents (*n* = 37) a version of this problem designed to make more accessible a refined, sixfold partition by drawing attention to the six card faces. The wording of this second version was identical to the first, except that the sides were numbered with relevant subscripts (red$_1$–red$_2$ for the first card, red$_3$–white$_1$ for the second card, and white$_2$–white$_3$ for the third card) and the last sentence read, "Given that the side showing is red, what is the probability that it is side Red$_1$ or Red$_2$?" We expected this new wording to facilitate a sixfold, refined partition by side: {red$_1$, red$_2$, red$_3$, white$_1$, white$_2$, white$_3$}. After the possibility of a white side {white$_1$, white$_2$, white$_3$} is edited from the sample space, it becomes apparent that the correct probability that the side is red$_1$ or red$_2$ is 2/3. Thus, we predicted that the original wording would give rise to a higher proportion of incorrect responses of 1/2 and that the revised wording would give rise to a higher proportion of correct responses of 2/3.

Results and Discussion

The results again revealed partition dependence consistent with our predictions (see Table 5). When presented with the modified wording, respondents were 10 times as likely to indicate the correct answer of 2/3 as participants presented with the original problem formulation (26% vs. 2.6%), $\chi^2(1, N = 76) = 9.18$, $p < .005$. Furthermore, most participants (59%) judged the probability to be 1/2 under the original problem wording, but only 24% did so under the alternative wording, $\chi^2(1, N = 76) = 9.35$, $p < .005$. Overall, our manipulation was successful in affecting the relative frequencies of responses implied by naive versus refined partitions under the partition–edit–count strategy, $\chi^2(1, N = 43) = 13.08$, $p < .001$.

Further support for the present account can be observed through an analysis of unconditional probabilities that derive from reliance on a naive versus refined partition. If a participant merely considered the probability that the other side (of a randomly selected card) is red without considering that the visible side is red, the symmetry of the problem would suggest a response of 1.5/3 = 1/2. Alternatively, some participants may have considered the conjunction of events (red face up and red underneath) and failed to edit. Such an analysis suggests that the red–red card had been selected, for which the unconditional probability is 1/3. Under the refined partition, however, there is no such ambiguity. If a participant considered the probability that a randomly selected side is red$_1$ or red$_2$ without considering that the target side is red, the probability would be 1/3. Indeed, 1/3 was the most common response other than 1/2 or 2/3 in both experimental conditions, and responses of 1/3 were more common under the modified problem wording than the original problem wording (35% vs. 13%), $\chi^2(1, N = 76) = 5.23$, $p < .05$.

Summary Analysis of Study 2

Numerical Responses

Taken together, Tables 3–5 indicate that the present account can explain the large majority of numerical responses provided by participants in Study 2. Application of the partition–edit–count strategy (to naive or refined partitions, with or without editing) can account for 87% of responses to the cancer drug problem, 93% of responses to the Mr. Smith problem, and 80% of responses to the three cards problem. Moreover, our experimental manipulation of the relative accessibility of naive versus refined partitions was apparently quite effective. Figures 1A and 1B summarize the proportion of participants by experimental condition that provided numerical responses predicted by application of the partition–edit–count strategy to refined and naive partitions, respectively. For all three substudies, responses consistent with use of a refined partition (i.e., the putative correct answer) were more common in the refined partition prime condition (the revised wording) than the naive partition prime condition (the original wording), and responses consistent with use of a naive partition (i.e., the dominant incorrect answer) were more common in the naive partition prime condition than in the refined partition prime condition.

Protocol Analysis

All respondents were asked to provide an explanation of their numerical responses. We hoped to obtain more direct evidence of partitioning and editing from these written protocols. A description of our procedure and instructions to coders are provided in the Appendix. The results of the protocol analysis are summarized in Table 6.

Over all studies, 60% of respondents explicitly invoked some form of partitioning when explaining their answer, 13% implicitly invoked partitioning but asserted that the conditioning information provided no news concerning the target event[7] (cf. Shimojo & Ichikawa, 1989), 11% relied on an explicitly computational ap-

---

[7] In the cancer drug problem, some participants reported that learning about the relative ranking of *A* and *C* told them nothing about the relative ranking of *A* and *B*, and therefore the probability is 1/2. In the Mr. Smith problem, many participants observed that the probability of a boy versus girl is independent for each child, so the probability of the other child being a boy or girl is 50–50. Both of these "no news" responses might be interpreted as implicitly expressing naive partition reasoning.
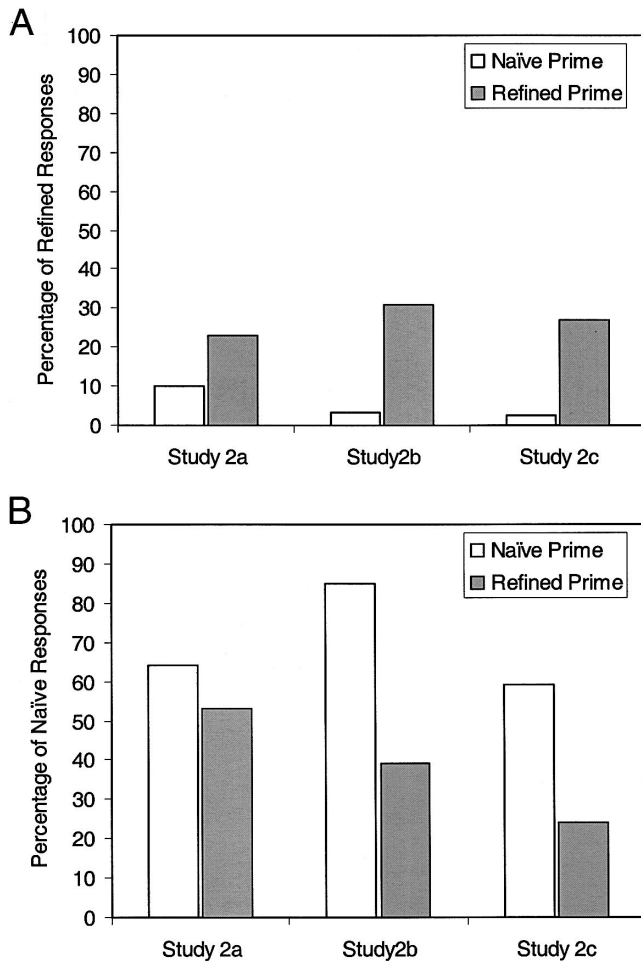
Figure 1. A: Refined partition–edit–count responses by experimental condition. Bars depict the percentage of numerical responses in each study that coincide with application of the partition–edit–count model to the corresponding refined partition. B: Naive partition–edit–count responses by experimental condition. Bars depict the percentage of numerical responses in each study that coincide with application of the partition–edit–count model to the corresponding naive partition.

proach, 4% used some other coherent strategy, and the remaining 12% gave explanations that were not interpretable or left the space blank. The proportion of participants who relied on partitioning did not vary significantly by condition for any of the substudies. It is striking that not a single participant who reported an explicit computation arrived at a correct answer.

The third-to-last and second-to-last columns of Table 6 show that we were successful in manipulating the relative accessibility of naive versus refined partitions through our variations in wording of each problem. For the cancer drug problem, participants who explicitly invoked a partition were much more likely to invoke the refined partition under the ranking formulation (A ranked ahead of B and C) than under the original formulation (A most effective), and they were more likely to invoke the naive partition under the original formulation, $\chi^2(1, N = 73) = 3.10$, $p = .078$. For the problem of Mr. Smith's children, participants who explicitly articulated a partition were much more likely to invoke the refined

partition under the revised formulation of the problem ("both children are boys") than under the original formulation ("the other child is a boy"), and they were far more likely to invoke a naive partition under the original wording, $\chi^2(1, N = 49) = 21.48$, $p < .0001$. In addition, 37% of participants who articulated a partition under the revised language of the Mr. Smith problem invoked a threefold partition that distinguished family composition but not birth order: {two boys; one boy and one girl; two girls}. For the three cards problem, participants who explicitly mentioned a partition were much more likely to invoke the refined partition under the revised (sides) formulation of the problem than under the original (cards) formulation, and they were more likely to invoke a naive partition under the original wording, $\chi^2(1, N = 52) = 8.10$, $p < .005$.

Table 7 shows that participants tended to arrive at answers that were consistent with application of the partition–edit–count strategy to the partitions that they articulated (see Column 5). Overall, 70% of participants in Studies 2A–2C who explicitly articulated a partition arrived at the numerical response predicted by application of the partition–edit–count strategy to that partition, and an additional 22% of these respondents arrived at the numerical response predicted by application of the partition–count strategy (with failure to edit) to that partition (see Column 6). Only 8% of these responses were inconsistent with either form of the partition–edit–count strategy.

Study 3: Partition Dependence and the Individuation Hypothesis

We have argued that people typically assess conditional probability by first subjectively partitioning the sample space into a set of elementary events that are treated as interchangeable and therefore equiprobable. We have argued further that mistakes made in answering probability puzzles can often be traced to inappropriate editing or use of a partition that is insufficiently refined, so that elementary events do not remain equiprobable in light of conditioning information.

Given this information, what factors influence the relative accessibility of alternative partitions? Brase et al. (1998) advanced the individuation hypothesis, according to which the "human mechanism for assessing relative frequencies is better designed for operating over whole objects than arbitrary parsings of them" (p. 9). The authors ran a set of studies using variations of Bar-Hillel and Falk's (1982) three cards in a hat problem and attributed the high frequency of "1/2" responses to a tendency to operate over representations of whole objects (i.e., partition the sample space into whole objects): "By whole object, we mean cohesive, bounded entities that move as a unit, independent of other surfaces. . . . The more closely a 'part' of an object conforms to this definition, the easier it should be to count" (p. 8).

To support their claim, Brase et al. (1998) presented respondents with isomorphs of the three cards problem in which they manipulated the wholeness of objects. For example, in one study, a first group of participants received a problem in which candy canes were sampled at random, with some that were lemon, some that were peppermint, and a third type that were lemon on one end and peppermint on the other end. Other participants received a version in which candy canes were sampled from jars, with one jar containing all lemon candy canes, one jar containing all pepper-

Table 6
Summary Protocol Analysis: Explanations Provided by Participants

Target event	Primed partition	% interpretable	% of interpretable solutions				% of partitions		
			Partition	No news	Comp	Other	Naive	Refined	Other
Experiment 2A: Cancer drugs									
A most effective	{A, B, C}	78	77	6	4	13	80	13	8
A ranked ahead of B, C	{ABC, ACB, BAC, BCA, CAB, CBA}	77	75	6	10	8	69	31	0
Experiment 2B: Mr. Smith's children									
Other child is B	{B, G}	92	40	47	9	4	73	18	9
Both children are Bs	{BB, BG, GB, GG}	89	84	0	16	0	9	54	37
Experiment 2C: Three cards in a hat									
Other side of card is R	{RR, RW, WW}	95	78		22	0	86	10	3
It is side R_1 or R_2	{R_1, R_2, R_3, W_1, W_2, W_3}	81	77		23	0	54	46	0

Note. *Primed partition* refers to the partition that the corresponding target event was designed to facilitate. *% interpretable* refers to the percentage of respondents who provided a coherent explanation of their response. The fourth through seventh columns list the percentage of participants who explained their answer by explicitly invoking a partition, indicating that the conditioning information provides no relevant information, using a computational strategy, and providing some other coherent explanation, respectively. The final three columns list the percentages of those participants who explicitly invoked a partition that relied on naive, refined, or other varieties of partition, respectively. In Experiment 2B, B = boy and G = girl. R = red; W = white; Comp = computational approach.

mint candy canes, and a third jar containing equal numbers of each flavor. Participants were more likely to arrive at the correct answer for the version described in terms of whole candy canes than for the version described in terms of candy canes that had different flavors on different ends.

We agree that in many situations the most accessible partition may correspond to the level of whole objects, as in Brase et al.'s (1998) studies. However, we argue that one's partition of the sample space is an ad hoc construction that is inherently malleable and is affected by whatever grouping happens to be made most accessible to the judge. Hence, unlike Brase et al., we predict that one can lead participants to segregate or integrate whole objects through subtle variations in wording of the three cards or any other problem, so that they can be induced to arrive at the correct or incorrect response. To wit, note that in Study 2C, we primed participants to segregate whole objects into parts of objects (cards into sides of cards), which helped them to arrive at the correct answer more often. Moreover, we assert that even when the correct answer can be derived from a ratio of whole objects, people can be led to invoke a coarser partition (and thereby arrive at an incorrect answer) through instructions that suggest that these objects be grouped.[8]

Method

To test the partition–edit–count model against the individuation hypothesis, we recruited entering MBA students at Duke University (*N* = 68) and asked them to answer the following question as part of a larger survey packet, in exchange for a donation to charity:

> Imagine that you are a venture capitalist in Palo Alto. You attend an industry luncheon at which the featured guests are pairs of managers from three new ventures. An insider has tipped you off to the fact that Company A consists of two strong managers, Company B consists of

a strong manager and a weak manager, and Company C consists of two weak managers.

> Company A: Strong–Strong
> Company B: Strong–Weak
> Company C: Weak–Weak

One group of participants (*n* = 33) was then asked a question designed to prime a naive (threefold) partition by company:

> You are seated next to Katherine, who is a manager from one of the three featured companies. By the end of lunch you are convinced that she is a strong manager. Given this assessment, what is your best estimate of the probability that the other manager from her company is also strong?

A second group of participants (*n* = 35) was asked a question designed to prime a refined (sixfold) partition by manager:

> You are seated next to Katherine, who is one of the six featured managers. By the end of lunch you are convinced that she is a strong manager. Given this assessment, what is your best estimate of the probability that she is one of the two strong managers from Company A?

Note that in Study 2C, the refined partition entailed sides of cards, which are part objects. However, in Study 3, the refined partition entails managers, which are certainly whole objects. Thus, Brase et al.'s (1998) individuation hypothesis seems to suggest that participants should have a relatively easy time answering this question correctly regardless of experimental condition. In contrast, the partition–edit–count model predicts that the company frame will prompt participants to invoke an unre-

---

[8] Indeed, the three cards in a hat problem is modeled after Bertrand's box paradox (see Nickerson, 1996, p. 417), which allows a correct analysis in terms of whole objects (gold and silver coins placed in three drawers of a box) but has nevertheless confounded puzzle enthusiasts for many decades.

Table 7
Summary of Protocol Analysis: Numerical Responses by Reported Partition

Reported partition	n	Response predicted by P-E-C	Response predicted by P-C	% indicating P-E-C response	% indicating P-C response	% indicating other response
			Experiment 2A: Cancer drugs			
{A, B, C}	57	1/2	1/3	76	21	3
{ABC, ACB, BAC, BCA, CAB, CBA}	16	2/3	1/3	75	6	19
			Experiment 2B: Mr. Smith's children			
{B, G}	20	1/2		98		2
{BB, BG, GB, GG}	19	1/3	1/4	62	31	7
			Experiment 2C: Three cards in a Hat			
{RR, RW, WW}	38	1/2	1/3	76	21	3
{R_1, R_2, R_3, W_1, W_2, W_3}	14	2/3	1/3	36	50	14

Note. P-E-C = partition–edit–count; P-C = partition–count; R = red; W = white. In Experiment 2B, B = boy and G = girl.

fined, threefold partition that facilitates an (incorrect) response of 1/2, whereas the manager frame will facilitate a refined, sixfold partition and a (correct) response of 2/3.

Results and Discussion

The results accord with our predictions (see Table 8). The median and modal judgment in the company (naive) frame was .50, whereas the median and modal judgment in the manager (refined) frame was .67 (Mann–Whitney, $p < .005$). Respondents were almost five times as likely to provide the correct answer of 2/3 in the manager frame compared with the company frame, $\chi^2(1, N = 68) = 9.95$, $p < .005$. Furthermore, respondents were about three times as likely to provide an answer of 1/2 in the company frame compared with the manager frame, $\chi^2(1, N = 68) = 13.35$, $p < .0005$. Overall, our manipulation was successful in affecting the relative frequencies of responses implied by naive versus refined partitions under the partition–edit–count strategy, $\chi^2(1, N = 46) = 14.94$, $p < .0001$. Thus, whole object representations do not necessarily facilitate correct answers, and the natural tendency to partition by whole objects can apparently be superceded by ad hoc conceptual groupings.

Study 4: Combining Partition–Edit–Count With Evaluative Strategies

We have argued that people naturally evaluate chance probabilities as a ratio of numbers of interchangeable events. Correct

Table 8
Results of Study 3

Condition	n	% responses			
		1/3	1/2	2/3	Other
Company (naive)	33	6	64	9	21
Manager (refined)	35	14	20	43	23

Note. The correct response is 2/3.

derivation of chance probabilities requires one to identify the focal outcome and the randomizing experiment that yielded it (cf. Bar-Hillel & Falk, 1982). In Studies 1A and 1B (the five-card problems) and Study 2C (three cards) the experiment entails shuffling, dealing, and flipping cards. In Study 2B (Mr. Smith) the experiment could be viewed as the genetic process that yields male and female children in roughly equal proportion. Studies 2A (cancer drug) and 3 (venture capitalist) extend this aleatory logic to epistemic uncertainty—if one has no way of distinguishing among pharmaceutical companies, one may naturally treat the labels attached to the first, second, and third ranked drugs as random; if one has no way of distinguishing among strong or weak managers, one may naturally treat the process by which one encounters a manager as random. Thus, in Studies 2A and 3, respondents apparently treated events that they could not distinguish the same way they might treat blind draws in a randomizing experiment and relied on the calculus of chance to derive subjective probabilities. Although applying the logic of aleatory uncertainty to epistemic uncertainty may raise thorny normative issues concerning the appropriate representation of outcomes and relevant randomizing experiment (see Bar-Hillel & Falk, 1982; Nickerson, 1996), our respondents seemed to find the partition–edit–count strategy intuitively compelling for such judgment under ignorance.

Study 4 extends the observation of partition dependence in conditional probability judgment from chance and ignorance to uncertainty, in which respondents are able to distinguish among elementary events using relevant knowledge or information. We predict that in these cases people will continue to rely on the partition–edit–count strategy but combine it with some form of evaluative assessment. Thus, if a variation in the description of a problem influences the edited partition that a person invokes, then judgments of conditional probability under uncertainty, like those under chance and ignorance, will be partition dependent. We designed Study 4 to investigate whether the partition–edit–count strategy is used in combination with an evaluative strategy when

participants are provided information that they can use to distinguish among events.

Method

We asked a new sample of entering MBA students at Duke University ($N = 71$) to complete the following item as part of a longer survey in exchange for a donation to charity:

> This coming Saturday and Sunday, the New York Mets will play two games in Los Angeles against the L.A. Dodgers. So far this season, the Mets have a win–loss record of 54 and 65; the Dodgers have a win–loss record of 65 and 54.

	W	L
Dodgers:	65	54
Mets:	54	65

Approximately half the participants ($n = 37$) were then asked the following question, designed to facilitate a naive partition, as in the Mr. Smith problem (Study 2B): "Assuming that the Dodgers win at least one game this weekend, what is your best estimate of the probability that the Dodgers win *both* games?"

The remaining participants ($n = 34$) were asked the same question in a way that was designed to facilitate a refined partition: "Assuming that the Dodgers do not *lose* both Saturday's game and Sunday's game, what is your best estimate of the probability that the Dodgers *win both* Saturday's game *and* Sunday's game?"

The present analysis suggests that the first wording will facilitate a twofold partition with no events edited out, so that many participants will anchor on 1/2, and the second wording should facilitate a fourfold partition with one event edited out, so that many participants will anchor on 1/3. Moreover, we expected both groups to adjust their probabilities upward because the Dodgers have a stronger record, are playing on their home field, and have won at least one game (the outcomes of the two games are presumably positively correlated). If participants anchor their judgments to

some extent on the output of the partition–edit–count strategy and then adjust using an evaluative strategy, and if they are more likely to invoke a refined partition in the refined partition prime condition than the naive partition prime condition (and vice versa), then (a) they should assign lower probabilities under the refined partition prime than under the naive partition prime, and (b) judged probabilities should vary somewhat from the values implied by the partition–edit–count strategy. Moreover, because the influence of the partition–edit–count strategy will be diluted by evaluative assessment and because there will be inevitable individual differences in how respondents evaluate the evidence, we expected the partition dependence observed in Study 4 to be less pronounced than the partition dependence we observed in previous studies.

Results and Discussion

The results of Study 4 accord with the notion that most participants relied on a combination of the partition–edit–count strategy and some form of evaluative assessment. First, in contrast to the results of the Mr. Smith problem (Study 2B), in which 93% of respondents reported probabilities of precisely 1/2, 1/3, or 1/4 (values corresponding to use of the naive partition, the refined edited partition, and the refined unedited partition, respectively), we obtained responses of precisely 1/2, 1/3 or 1/4 from only 35% of participants in the naive partition prime condition and 22% in the refined partition prime condition, and these proportions do not differ significantly, $\chi^2(1, N = 71) = 1.64$, *ns*. Thus, it seems that a large majority of participants relied on some criterion in addition to (or other than) the partition–edit–count strategy. Second, inspection of the cumulative distribution of responses by experimental condition, displayed in Figure 2, confirms that judged probabilities were generally lower (i.e., greater mass is to the left) under the language designed to make the refined partition more accessible (anchor of 1/3) than under the language designed to make the
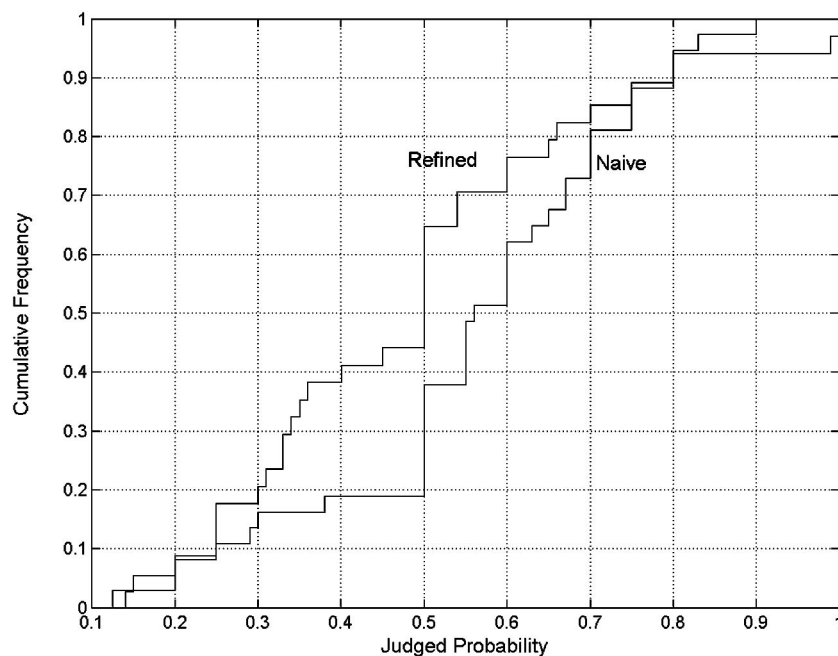
Figure 2. Results of Study 4. Plot of cumulative distribution of judged probabilities in naive partition prime (assumed anchor = 1/2) and refined partition prime (assumed anchor = 1/3) conditions.

naive partition more accessible (anchor of 1/2), and the difference in distributions is statistically significant (Kolmogorov–Smirnov $z = 1.38$, $p < .05$). Also, we note that the distribution of responses is consistent with our interpretation that participants in the refined partition prime condition were more likely to adjust their responses upward from 1/3 or 1/4 (values implied by use of the refined edited and unedited partition, respectively) and participants in the naive partition prime condition were more likely to adjust their responses upward from 1/2 (the value implied by use of the naive partition). For instance, 35% of participants in the refined partition prime condition reported probabilities in the range $.25 \leq p < .50$, whereas only 11% did so in the naive partition prime condition; 44% of participants in the refined partition prime condition reported probabilities in the range $.50 \leq p \leq .75$, whereas 70% did so in the naive partition prime condition, $\chi^2(1, N = 57) = 6.81$, $p < .01$. Thus, it seems that judgments of conditional probability under uncertainty were biased in the direction of values implied by use of the partition–edit–count strategy.

General Discussion

In this article we have argued that when people evaluate conditional probabilities they commonly invoke a strategy in which they subjectively partition the sample space into n elementary possibilities, edit out possibilities that can be eliminated on the basis of conditioning information, then count the remaining possibilities and report probabilities as the ratio of the number of focal events to the total number of events. A first study (1A), which used a modified version of the famous Monty Hall problem, provides evidence of partition editing and its consequences. Judged probabilities were influenced by an experimental manipulation that did not affect normative probabilities but apparently induced participants to eliminate particular events from their subjective partition. A follow-up study (1B) demonstrates that this tendency to edit is stronger when the conditioning information is communicated in a concrete (visual) way rather than an abstract (verbal) way. A second set of studies demonstrates that rewording well-known probability puzzles in ways that increase the accessibility of more refined partitions can reduce the frequency of errors committed by participants. An informal protocol analysis provides additional support for our interpretation that responses are driven by use of the partition–edit–count strategy. A third study demonstrates that partitions may be constructed on the basis of an ad hoc grouping of items (e.g., firms represented by two managers) and are not necessarily driven by the identification of whole objects, as suggested by previous research (Brase et al., 1998). Finally, a fourth study extends the present analysis from the domains of chance and ignorance, in which people do not distinguish among elementary events, to the domain of uncertainty, in which people do distinguish among events. In this case people apparently rely on the partition–edit–count strategy in combination with some form of evaluative assessment. We close with a discussion of related approaches, determinants of partitioning and editing, more formal modeling of the partition–edit–count strategy, and implications of the present work for teaching probabilistic reasoning.

Related Approaches

A number of authors have presented accounts of probabilistic intuitions that are related to the partition–edit–count model and

have applied them to some of the chance illusions that we have presented here. Falk (1992) attributed dominant responses to the three prisoners/Monty Hall problem to two primary intuitions: the uniformity intuition, which refers to the belief that probabilities are uniformly distributed, and the no news, no change intuition, which refers to the belief that the conditioning information does not change the relative probability of the remaining alternatives. However, she provided no empirical studies to verify these assertions.

Shimojo and Ichikawa (1989) also studied intuitive reasoning about probabilities, and, like Falk (1992), they restricted their investigation to the three prisoners problem. Using protocol analysis, these authors identified three primary strategies that participants use to solve this problem. The authors' constant ratio theorem ("when one alternative is eliminated, the ratio of probabilities for the remaining alternatives is the same as the ratio of prior probabilities for them"; p. 7) could be viewed as a single-step combination of editing and counting of an unrefined partition.

Finally and most notable, Johnson-Laird and colleagues (1999) suggested that people construct mental models of possibilities that might be true (the truth principle), assume that these mental models are equally likely to be true (the equiprobability principle), and then compute probability as the proportion of mental models in which a target event occurs (the proportionality principle). These authors focused primarily on logical propositions with sentential connectives. For instance, consider one of their simplest examples, "There is a box in which there is at least a red marble, or else there is a green marble and there is a blue marble, but not all three marbles . . . What is the probability [that] there is a red marble and a blue marble?" (p. 74). Johnson-Laird et al. (1999) argued that people will spontaneously represent the two mental models {red; green and blue} and treat each of these possibilities as equally likely, so that the judged probability of both a red marble and a blue marble should be 0%. They argued that a fully explicit model of the premises should take into account that where it is true that there is a red marble, there are also three distinct ways it can be false that there is both a green marble and a red marble, which can be represented by the four mental models {red and green and not blue; red and not green and blue; red and not green and not blue; not red and green and blue}, yielding an unbiased probability of 25%.

Although the partition–edit–count model is similar to previous interpretations of conditional probability judgment, there are some important theoretical distinctions. First, the present account emphasizes that partitions and edits are subjective and can be influenced by the language of a probability query or the modality in which information is presented (e.g., visual vs. verbal), whereas foregoing approaches assumed a canonical representation of the sample space for a given problem. Indeed, mental model theory has been criticized on the grounds that it is a deductive system that assumes a fixed representation without much discussion of where that representation comes from (Lagnado & Sloman, in press). Second, unlike foregoing models, the present account explicitly segregates editing from partitioning, which brings into sharper focus the importance of editing (or lack thereof) in probability assessment and subsequent choice. Study 1A and the analysis of unconditional probabilities in Study 2 demonstrate the importance of the editing process, and Study 1B suggests that the accessibility of edits is influenced by how concretely conditioning information is communicated. Third, mental model theory assumes that only

true statements are represented as mental models, so that the set of mental models considered is not necessarily an exhaustive partition of the sample space. Indeed, the main emphasis of the Johnson-Laird et al. (1999) article is to trace biases in probability assessment to a tendency to *not* explicitly represent false statements, as in the example provided in the preceding paragraph. Although we do not take issue with Johnson-Laird et al.'s analysis of their problems, the main emphasis of the present work is on how conditional probability judgments depend on information that *is* represented (the partition and its associated edits). Fourth, although Johnson-Laird et al. (1999) and Shimojo and Ichikawa (1989) relaxed their equiprobability assumptions to allow for chance probabilities that are unequal across mental models or events, they did not provide insight into the origin of these unequal probabilities, and they assumed that these exogenous probabilities were known with precision (see Lagnado & Sloman, in press). The present account allows for adjustment of conditional probabilities on the basis of respondents' relevant knowledge or information concerning how elementary events differ (Study 4); we return to this point in greater detail shortly.

In addition to these theoretical distinctions, the present work differs from previous studies in several important methodological respects. First, unlike Johnson-Laird et al. (1999) and Shimojo and Ichikawa (1989), we experimentally manipulate the relative accessibility of alternative representations of sample spaces (partitions and edits) and observe concomitant shifts in assessed probability. Moreover, by explicitly perturbing each stage of the process, we are able to test more directly the association between participants' representations of the sample space and their resulting judgments. Second, unlike Johnson-Laird et al. (1999), we bolster support for our account with an analysis of written protocols that allows us to trace the association between the problem description and the reasoning invoked by participants, as well as the association between the reasoning invoked and assessed probability (Study 2). Third, we use a more diverse set of problems than previous researchers and extend our investigation beyond the domain of chance events to the domain of uncertainty, in which evaluative assessment may also play a role. In addition, we have included some items in which participants must act on their beliefs and in which there are real monetary consequences (Studies 1A and 1B).

Determinants of Partitioning and Editing

The phenomenon of partition dependence has previously been observed for simple (unconditional) probabilities in a number of other domains (Fox & Clemen, 2004; Fox & Rottenstreich, 2003; See, Fox, & Rottenstreich, 2004). In these cases, people may anchor their beliefs on a quasi-Bayesian "ignorance prior" probability derived from a consideration of the possible events that might obtain (i.e., a subjective partition of the sample space) and adjust according to an assessment of the distinguishing features of those events. This may be akin to use of a partition–count strategy (note that for unconditional probabilities there is no need to edit).

Taken together, the present article and previous investigations of partition dependence have identified a number of varieties of partition that people naturally invoke. Fox and Rottenstreich (2003) argued that the default partition is a twofold, "case" partition defined by the target event and its complement. For instance, when judging the likelihood of rain next Sunday, one might

consider the partition {it will rain Sunday; it will not rain Sunday}. The authors distinguished this from an *n*-fold, "class" partition of events that can be viewed in some respect as interchangeable. For example, when judging the likelihood that a particular horse will win a race, one might consider the partition {Horse 1 will win; Horse 2 will win; . . . Horse *n* will win}. The present article distinguishes different varieties of class partition (naive vs. refined).

What factors influence partitions that people invoke? First, we suggest that the default partition may be a (twofold) case partition because it is always easy to construct: The target event either will or will not occur. Indeed, when Fischhoff and Bruine de Bruin (1999) presented college students with unfamiliar questions such as, "What are the chances that you will get cancer by age 40?" many students responded .50, reasoning that the event could either happen or not happen. Second, participants may adopt class partitions when they are explicitly asked to judge probabilities of a particular set of exclusive and exhaustive events, as is common in decision analysis. Fox and Clemen (2004) found that participants with training in decision analysis were biased toward assigning probabilities of 1/*n* to each of the *n* events into which the sample space had been explicitly partitioned, and this bias was stronger (i.e., there was apparently less overall adjustment) for less familiar domains. Third, the language of a probability query may influence the relative accessibility of alternative partitions. Fox and Rottenstreich (2003) demonstrated that case partitions were more accessible when queries highlighted the target event, whereas class partitions were more accessible when queries highlighted a class of interchangeable events. For instance, participants who were asked to judge the probability that "the temperature on Sunday will be higher than every other day next week" were more likely to respond with the ignorance prior probability implied by the case partition (1/2) and less likely to respond with the ignorance prior probability implied by the class partition (1/7) than participants who were asked to judge the probability that "next week, the highest temperature of the week will occur on Sunday." Likewise, the present article demonstrates that the relative accessibility of naive versus refined class partitions can be influenced by the description of the initial conditions and conditioning information. Fourth, different partitions may be suggested by topological features of the relevant sample space. For instance, See et al. (2004) showed that a class partition may be suggested by the number of distinct levels that a target attribute takes on. In some studies, participants observed the frequency of colored shapes that flashed on a computer screen. When subsequently asked to judge the probability of a particular attribute (e.g., black objects or triangles), they were biased toward the ignorance prior probability defined by the number of values that the target attribute could take on (black was one of two possible colors; triangles were one of four possible shapes), even when these attributes appeared with the same frequency. This bias was more pronounced (i.e., probabilities were closer to the corresponding ignorance prior) when participants had a less extensive opportunity to learn.

What factors influence editing operations? We surmise that conditioning information can influence judgment under uncertainty in one of two ways. First, conditioning information may lead people to modify their evaluation of the relative strength of evidence for events under consideration. For instance, learning that Ames received warmer letters of reference than Clark can lead one

to update one's prior probability of 1/3 to a more optimistic assessment of Ames's chances relative to the other candidates. Second, conditioning information may clearly rule out the possibility that some events will occur and thereby trigger an editing operation. Thus, learning that Ames is a stronger candidate than Clark leads one to edit the event *Clark is strongest* from the initial partition. Moreover, the accessibility of an editing operation may be influenced by how concretely conditioning information is presented or represented—apparently participants found editing to be somewhat more natural in Study 1B when nonace cards were visually revealed than when they were verbally identified.

We suspect that reliance on some form of the partition–edit–count strategy is ubiquitous in conditional probability assessment. In many situations, use of this strategy may be difficult to detect because people may invoke a trivial twofold case partition, and conditioning information therefore does not eliminate either of these events from consideration but rather affects the perceived relative strength of evidence for these possibilities. For instance, in assessing the likelihood that Linda is a bank teller given a particular description of Linda, one may begin with an initial case partition {Linda is a bank teller; Linda is not a bank teller} and an ignorance prior probability of .50, then adjust this probability according to the relative similarity of Linda to one's stereotype of bank tellers compared with other possible professions that come to mind. On the other hand, class partitions and editing operations may be especially accessible in situations of chance or ignorance because these situations provide little opportunity to distinguish among events using evaluative strategies, as the opening Ames–Boyd–Clark example illustrates.

Incorporating Partition–Edit–Count Into Support Theory

To model partition dependence, Fox and Rottenstreich (2003) advanced a formal refinement of support theory (Rottenstreich & Tversky, 1997; Tversky & Koehler, 1994) that incorporates both the ignorance prior and the balance of evidential support. In support theory, each hypothesis H is associated with a nonnegative support value $s(H)$ that is interpreted as the strength of evidence for this hypothesis. The probability $P(H_f, H_a)$ that the focal hypothesis H_f (e.g., the Dodgers win both games) holds rather than the alternative hypothesis H_a (e.g., the Dodgers fail to win both games) is given by $P(H_f, H_a) = s(H_f)/[s(H_f) + s(H_a)]$. Rewriting this formula in an odds metric yields $R(H_f, H_a) \equiv P(H_f, H_a)/[1 - P(H_f, H_a)] = s(H_f)/s(H_a)$. Support arising from the ignorance prior can be segregated from support generated by evaluative assessment if we rewrite the odds form as $R(H_f, H_a) = [n_f / n_a]^{1-\lambda}[s*(H_f)/s*(H_a)]^\lambda$, with $0 \leq \lambda \leq 1$. Here n_f and n_a are the number of elements in the subjective partition that correspond to the focal and alternative hypotheses, respectively, $s*(H_f)$ and $s*(H_a)$ quantify support generated through evaluative assessment, and λ indicates the relative contribution of these two sources of support. As λ approaches 0, judgments converge to the ignorance prior; when $\lambda = 1$, judgments are based entirely on evaluative assessment.

The present account of conditional probability judgment can be incorporated readily into the ignorance prior model if we assume that the ignorance prior is derived through the partition–edit–count strategy. In the case of judgment under chance and ignorance (Studies 1–3), there is no adjustment (i.e., $\lambda = 0$). In the case of

judgment under uncertainty (Study 4), there is adjustment (i.e., $\lambda > 0$). In both cases, the ignorance prior is assumed to vary with the description of the problem. Hence, when participants are asked to judge the probability that the Dodgers win both games, assuming that they win at least one game, participants may invoke a simple twofold partition in which the Dodgers either do or do not win both games {win both, do not win both}, with $n_f = 1$ and $n_a = 1$. When asked to judge the probability that the Dodgers win both games, given that they have not lost both Saturday's and Sunday's games, participants may be more likely to invoke an (edited) refined partition in terms of Saturday's and Sunday's results {win–win, win–lose, lose–win} with $n_f = 1$ and $n_a = 2$. The resulting ignorance priors would be adjusted according to the balance of evidence for the Dodgers winning both games rather than failing to win both games, which may reflect a consideration of the relative records of the teams, the information that the Dodgers beat the Mets at least once in the weekend series, and/or any outside knowledge that the participants bring with them concerning baseball and these teams.[9] Thus, the partition–edit–count strategy may provide new insights into the origin of the ignorance prior in conditional probability judgment and draw attention to new sources of bias: the choice of a naive versus refined partition or the tendency to edit. Future work is needed to more fully flesh out the implications of this model and measure its parameters.

One might wonder whether the refined partition prime not only enhances the accessibility of a refined partition but also cues people to mentally unpack events more thoroughly and evaluate evidence for a richer set of possibilities. For instance, it could be that "assuming the Dodgers do not lose both Saturday's and Sunday's games . . . what is the probability that the Dodgers win both games?" facilitates greater attention to scenarios in which the Dodgers fail to win both games (e.g., ways they could win Saturday and lose Sunday or lose Saturday and win Sunday) than does "assuming the Dodgers win at least one game . . . what is the probability that the Dodgers win both games?" Further research is required to understand the extent to which varying the language of the probability query facilitates alternative partitioning versus unpacking of evidence in judgment under uncertainty, as in the Dodgers–Mets example of Study 4. Indeed, these accounts are not mutually exclusive. However, we surmise that the partition–edit–count model provides a more parsimonious account of conditional probability judgment under chance and ignorance, as in Studies 1–3, because (a) evaluative assessment is not available in these cases, and (b) the partition–edit–count model uniquely predicts the high proportion of judgments that precisely equal ignorance priors implied by naive versus refined partitions.

Implications for Teaching Probabilistic Reasoning

In this article we have found evidence of the partition–edit–count strategy among participants with a range of sophistication. Participants in Study 1 were members of a university community (mostly undergraduate students), participants in Studies 3 and 4 were entering MBA students (university graduates returning to

---

[9] For a discussion of how conditioning information affects the evaluation of evidence in support theory, see Tversky and Koehler (1994, pp. 550–551).

school after some work experience), and participants in Study 2 were MBA students who had recently received training in probability and statistics. It is surprising that the vast majority of participants in the latter group who explained their answer with an identifiable strategy invoked a partition rather than an explicitly computational approach. Moreover, despite the fact that participants could have solved all three puzzles computationally by invoking Bayes's theorem or the definition of conditional probability, a very small proportion of these respondents seemed to attempt a computational answer, and none of the participants who explicitly invoked a formula arrived at the correct solution.

Apparently, even sophisticated participants resonate with a conception of probability as a ratio of interchangeable events rather than a more abstract computation. Although the partition–edit–count strategy can provide a correct answer to many probability problems, our studies suggest that people often fall prey to two common errors when taking such an approach: (a) They invoke a partition that is insufficiently refined so that the elements are no longer equiprobable in light of the conditioning information, and/or (b) they fail to edit appropriately the cases that can be eliminated.

With these observations in mind, we suggest that introductory courses in probability might begin by acknowledging this intuitive predilection. For instance, such courses might give some attention to the questions of what constitutes an appropriate partition in light of the conditioning information and the experiment that generates it. Indeed, Nisbett, Krantz, Jepson, and Kunda (1983) observed that greater clarity of the sample space and sampling process facilitates the use of strategies based on formal statistical principles (see also Nisbett, Fong, Lehman, & Cheng, 1987). After students are proficient in this intuitive approach to conditional probability, they may be more receptive to a more flexible, computational approach. Indeed, Sedlmeier and Gigerenzer (2001) reported more success teaching Bayesian reasoning to students when problems were presented in simple frequency format rather than probability format. We surmise that simple frequency formats facilitate appropriate use of the partition–edit–count strategy.

References

Bar-Hillel, M., & Falk, R. (1982). Some teasers concerning conditional probabilities. *Cognition, 11,* 109–122.

Ben-Zeev, T., Stibel, J., Dennis, M., & Sloman, S. (2003). *Increasing working memory demands improves probabilistic choice but not judgment.* Unpublished manuscript, Department of Psychology, San Francisco State University.

Brase, G. L., Cosmides, L., & Tooby, J. (1998). Individuation, counting and statistical inference: The role of frequency and whole-object representations in judgment under uncertainty. *Journal of Experimental Psychology: General, 127,* 3–21.

de Finetti, B. (1931). Sul significato soggetivo della probabilità [On the subjective meaning of probability]. *Fundamenta Mathematical, 17,* 298–329.

de Finetti, B. (1937). La prevision: Ses lois logiques, ses sources subjectives [Forecasting: Its logical laws, its subjective sources]. *Annales de l'Institut Henri Poincaré, 7,* 1–68.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis.* Cambridge, MA: MIT Press.

Falk, R. (1992). A closer look at the probabilities of the notorious three prisoners. *Cognition, 43,* 197–223.

Falk, R., & Wilkening, F. (1998). Children's construction of fair chances: Adjusting probabilities. *Developmental Psychology, 34,* 1340–1357.

Fischhoff, B., & Bruine de Bruin, W. (1999). Fifty–fifty = 50%? *Journal of Behavioral Decision Making, 12,* 149–163.

Fox, C. R., & Clemen, R. T. (2004). *Partition dependence in subjective probability assessment: Evidence for a bias toward the ignorance prior.* Unpublished manuscript, Anderson School of Business, University of California, Los Angeles.

Fox, C. R., & Rottenstreich, Y. (2003). Partition priming in judgment under uncertainty. *Psychological Science, 13,* 195–200.

Franco-Watkins, A. M., Derks, P. L., & Dougherty, M. R. P. (2003). Reasoning in the Monty Hall problem: Examining choice behaviour and probability judgements. *Thinking and Reasoning, 9,* 67–90.

Friedman, D. (1998). Monty Hall's three doors: Construction and deconstruction of a choice anomaly. *American Economic Review, 88,* 933–946.

Gardner, M. (1961). *The 2nd Scientific American book of mathematical puzzles and diversions.* New York: Simon & Schuster.

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment.* Cambridge, England: Cambridge University Press.

Gilovich, T., Medvec, V. H., & Chen, S. (1995). Commission, omission, and dissonance reduction: Coping with regret in the "Monty Hall" problem. *Personality and Social Psychology Bulletin, 21,* 182–190.

Girotto, V., & Gonzalez, M. (2003). Probabilistic reasoning and combinatorial analysis. In V. Girotto & P. N. Johnson-Laird (Eds.), *The shape of reason: Essays in honour of Paolo Legrenzi.* London, England: Psychology Press.

Granberg, D., & Brown, T. A. (1995). The Monty Hall dilemma. *Personality and Social Psychology Bulletin, 21,* 711–721.

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M. S., & Caverni, J.-P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review, 106,* 62–88.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge, England: Cambridge University Press.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology, 17,* 248–294.

Kotovsky, K., & Simon, H. A. (1990). What makes some problems really hard: Explorations in the problem space of difficulty. *Cognitive Psychology, 22,* 143–183.

Kraus, S., & Wang, X. T. (2003). The psychology of the Monty Hall problem: Discovering psychological mechanisms for solving a tenacious brain teaser. *Journal of Experimental Psychology: General, 132,* 3–22.

Kyburg, H. E., & Smokler, H. E. (1980). *Studies in subjective probability.* New York: Kriegler.

Lagnado, D. A., & Sloman, S. A. (in press). Inside and outside probability judgment. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making.* London: Blackwell.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32,* 311–328.

Morgan, J. P., Chaganty, N. R., Dahiya, R. C., & Doviak, M. J. (1991). Let's make a deal: The player's dilemma. *American Statistician, 45,* 284–287.

Nickerson, R. (1996). Ambiguities and unstated assumptions in probabilistic reasoning. *Psychological Bulletin, 120,* 410–433.

Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science, 238*(4827), 625–631.

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

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

Rottenstreich, Y., & Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review, 104,* 406–415.

Sedlmeier, P., & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General, 130,* 380–400.

See, K. E., Fox, C. R., & Rottenstreich, Y. (2004). *Attribute-cued partitioning and learning in judgment under uncertainty*. Unpublished manuscript, Wharton School of Management.

Shimojo, S., & Ichikawa, S. (1989). Intuitive reasoning about probability: Theoretical and experimental analyses of the "problem of three prisoners." *Cognition, 32,* 1–24.

Tor, A., & Bazerman, M. H. (2003). Focusing failures in competitive environments: Explaining decision errors in the Monty Hall game, the Acquiring a Company problem, and multiparty ultimatums. *Journal of Behavioral Decision Making, 16,* 353–374.

Tversky, A., & Kahneman, D. (1983). Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90,* 293–315.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business, 59,* S251–S278.

Tversky, A., & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review, 101,* 547–567.

vos Savant, M. (1990). Ask Marilyn. *Parade Magazine,* p. 15.

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology and reasoning: Structure and content.* London, England: Batsford.

Appendix

Protocol Analysis Procedure

All participants in Study 2 were asked to provide a written explanation of their numerical responses. Because these protocols were informal and not directed in real time (Ericsson & Simon, 1993), the results of this analysis should be regarded as preliminary. Two judges who were blind to our hypotheses and also blind to the experimental conditions coded the protocols in two phases (coding instructions are provided below). In the first phase, the coders were asked to classify the primary rationale articulated by respondents into one of four or five categories, depending on the item, one of which was explicit use of partitioning. Initial ratings of the judges in this first phase agreed in 70% of the cases for the cancer drug problem, 83% for Mr. Smith's children, and 81% for the three cards problem. Disagreements were resolved by discussion. In the second phase, coders were asked to reexamine the protocols classified as partitioning and identify the number of elementary events in the partition invoked by the respondent (e.g., threefold, sixfold). Initial rates of agreement were 95% for the cancer drug problem, 85% for Mr. Smith's children, and 81% for the three cards problem. Disagreements were again resolved by discussion.

Phase 1

Please do your best to discern the *primary* mode of reasoning that each participant invokes in explaining his or her response. In some cases it may appear that the participant is invoking multiple explanations or modes of reasoning. In such cases please do your best to identify which form of reasoning is the primary mode from which their numerical answer seems to follow. If you think that there is a second mode of reasoning that is plausible, please write this down in parentheses as it will make your job easier of resolving disagreements with the other coder later.

Pharmaceuticals

1. *Explicit partitioning*. This category applies when the respondent identifies a set of (exclusive and exhaustive) possibilities that are treated as equiprobable. Explicit partitioning can be exhibited either through a verbal statement (e.g., "It could be A or B but not C") or a symbolic representation (e.g., writing A B C, then crossing out C or writing ABC, ACB, BAC, BCA, CAB, CBA, then crossing out some of these elements).

2. *Independence/no news*. This category applies when respondents report that learning about the relative ranking of A and C tells us nothing about the relative rank of A and B (and therefore the probability of A is unchanged in light of this information).

3. *Computation*. This category applies when respondents derive their response from some kind of calculation using an equation, or the explicit multiplication, division, addition, and/or subtraction of fractions or decimals. Computation might entail plugging values into an equation (e.g., Bayes's theorem or the definition of conditional probability) or multiplying through branches of a decision tree.

4. *Other coherent*. This category applies when respondents write some kind of coherent explanation that does not fit any of the above categories. By *coherent* we mean that the coder can see how the participant's answer follows from his or her reasoning.

5. *Uninterpretable/no explanation*. This category applies when respondents provide an incoherent attempt at an explanation, an explanation for which the reasoning is highly ambiguous, or no explanation at all.

Mr. Smith's Children

1. *Explicit partitioning*. This category applies when the respondent identifies a set of (exclusive and exhaustive) possibilities that are treated as equiprobable. Explicit partitioning can be exhibited either through a verbal statement (e.g., "The child is either a boy or girl") or a symbolic representation (e.g., writing B, G or writing BB, BG, GB, GG, then crossing out GG).

2. *Independence/no news*. This category applies when respondents indicate that the sexes of two children (or their probabilities) are independent.

3. *Computation*. This category applies when respondents derive their response from some kind of calculation using an equation, or the explicit multiplication, division, addition, and/or subtraction of fractions or decimals. Computation might entail plugging values into an equation (e.g., Bayes's theorem or the definition of conditional probability), or multiplying through branches of a decision tree.

4. *Other coherent*. This category applies when respondents write some kind of coherent explanation that does not fit any of the above categories. By *coherent* we mean that the coder can see how the participant's answer follows from his or her reasoning.

(Appendix continues)

5. *Uninterpretable/no explanation*. This category applies when respondents provide an incoherent attempt at an explanation, an explanation for which the reasoning is highly ambiguous, or no explanation at all.

Three Cards in a Hat

1. *Explicit partitioning*. This category applies when the respondent identifies a set of (exclusive and exhaustive) possibilities that are treated as equiprobable. Explicit partitioning can be exhibited either through a verbal statement (e.g., "It could only be the Red–White card or the Red–Red card") or a symbolic representation (e.g., writing down RR, RW, WW, then crossing out WW).

2. *Computation*. This category applies when respondents derive their response from some kind of calculation using an equation, or the explicit multiplication, division, addition, and/or subtraction of fractions or decimals. Computation might entail plugging values into an equation (e.g., Bayes's theorem or the definition of conditional probability), or multiplying through branches of a decision tree.

3. *Other coherent*. This category applies when respondents write some kind of coherent explanation that does not fit any of the above categories. By *coherent* we mean that the coder can see how the participant's answer follows from his or her reasoning.

4. *Uninterpretable/no explanation*. This category applies when respondents provide an incoherent attempt at an explanation, an explanation for which the reasoning is highly ambiguous, or no explanation at all.

Phase 2

Pharmaceuticals

There are three possible categories of partition:

1. *Threefold*. This category applies when a respondent identifies three elementary events (corresponding to which drug is "most effective"), then edits one out on the basis of the information provided. Score as threefold partition even if the respondent fails to edit or if the respondent edits before identifying all three elementary events (so that two remain—e.g., "either A or B is most effective").

2. *Sixfold*. This category applies when a respondent identifies six elementary events (corresponding to possible rankings of the drugs; e.g., ABC, BCA, CAB), then edits three possible orderings out on the basis of the information provided. Score as sixfold partition even if the respondent fails to edit or if the respondent edits before identifying all six elementary events (e.g., if the respondent lists only the three orderings that remain: ABC, ACB, BAC).

3. *Other*. This category applies when a respondent identifies a different variety of partition that does not fit into one of the two categories above.

Mr. Smith's Children

There are four possible categories of partition:

1. *Twofold*. This category applies when a respondent identifies two elementary events (corresponding to the sex of the "other" child).

2. *Threefold*. This category applies when a respondent identifies three elementary events (corresponding to possible sexes of two children; i.e., two boys, one boy and one girl, two girls) then edits one event out (two girls) on the basis of the information provided. Score as threefold partition even if the respondent fails to edit or if the respondent edits before identifying the three elementary events (e.g., if the respondent says Mr. Smith either has one boy and one girl or two boys).

3. *Fourfold*. This category applies when a respondent identifies four elementary events (corresponding to possible sexes of two children by birth order; i.e., BB, BG, GB, GG), then edits one event out (GG) on the basis of the information provided. Score as fourfold partition even if the respondent fails to edit or if the respondent edits before identifying the four elementary events (e.g., if the respondent lists BB, BG, GB).

4. *Other*. This category applies when a respondent identifies a different variety of partition that does not fit into one of the three categories above.

Three Cards in a Hat

There are three possible categories of partition:

1. *Threefold*. This category applies when a respondent identifies three elementary events (corresponding to each of the three cards), then edits one out on the basis of the information provided. Score as threefold partition even if the respondent fails to edit or if the respondent edits before identifying the three elementary events (so that two remain—e.g., "either RR or RW").

2. *Sixfold*. This category applies when a respondent identifies six elementary events (corresponding to the six sides of cards), then edits three sides on the basis of the information provided. Score as sixfold partition even if the respondent fails to edit or if the respondent edits before identifying the six elementary events (e.g., if the respondent lists only R1, R2, R3).

3. *Other*. This category applies when a respondent identifies a different variety of partition that does not fit into one of the two categories above.