Ambiguous probabilities and the paradoxes of expected utility

Chapter written for G. Wright and P. Ayton (Eds.) *Subjective probability.* Chichester, Sussex: Wiley, 1994.

Jonathan Baron Department of Psychology, University of Pennsylvania

Deborah Frisch Department of Psychology, University of Oregon]

1 Introduction

Recently the ambiguity effect (Ellsberg, 1961) has received a great deal of attention from psychologists and philosophers interested in decision theory (Einhorn & Hogarth, 1985; Frisch & Baron, 1988; Grdenfors & Sahlin, 1982; Heath & Tversky, 1991). The original ambiguity effect was the finding that people often prefer to bet on gambles with a known chance of winning as opposed to those where the chance of winning is unknown. For example, consider the following two gambles:

> Gamble 1: A marble will be drawn from an urn containing 50 black marbles and 50 white marbles. You win $100 if the marble is black. (Or, you can pick a color and you win if that color is drawn.)
>
> Gamble 2: An urn contains 100 marbles. Between 0 and 100 are black and the rest are white. A marble will be selected at random from the urn. You win $100 if the marble is black. (Or, you can pick a color and you win if that color is drawn.)

From the perspective of expected-utility theory, as we shall explain, these two gambles are equivalent. There is no reason to think that black is more or less likely than white in either case, and there is no other possible outcome. It therefore makes sense to think that the probability of winning is 1/2 in either case. Nonetheless, many people prefer Gamble 1. Ellsberg used the term *ambiguity* for the kind of unknown risk in Gamble 2. A situation in which the 'probability is unknown' is called *ambiguous*.

In principle, you can make money from someone who dislikes ambiguous bets (Camerer and Weber, 1992, p. 359). You can remove the ambiguity from Gamble 2 by flipping a coin in order to decide which color wins (Raiffa, 1961): the chance of winning is clearly 50% in this case. An ambiguity averter holding a ticket on 'black in Gamble 2' will therefore pay you to trade it for 'black if heads and white if tails.' Then flip the coin. If it is heads, do nothing, and you have been paid to return the person to her original state. If it is tails, get her to trade her bet on white for a bet on black. (Surely she is indifferent between these.) Again, she has paid you to get her back where she started.

Although this particular con game has apparently not been tried, Tversky and Kahneman asked subjects about the following game: 'Two boxes each contain red and green marbles. A marble is drawn from each box; if their colors match, you win \$60. In game A, both boxes have 50% red marbles and 50% green marbles. ... In game C, both boxes have the same composition of red and green marbles, but the composition is unknown' (cited by Camerer & Weber, 1992, p. 359). Most subjects preferred to play game A, but the chance of winning is higher in C. The decision rules that people follow thus fail to maximize their winnings in the long run. This fact suggests that aversion to ambiguity is an error. We shall examine this suggestion.

In a three-color version of the Ellsberg paradox, an urn contains 90 balls. Thirty of them are red, and 60 of them are either black or yellow - we do not know which. A ball is to be drawn from the urn, and we can win some money, depending on which ball is drawn and which option we take, as shown in the following table:

	30 red balls	60 balls	
		black	yellow
Option X	\$100	\$0	\$0
Option Y	\$0	\$100	\$0

Most subjects lean toward option X. They 'know' that they have a 1/3 chance of winning in this case (30 out of 90 balls). They do not like option Y because they feel that they do not even know what the 'real probability' of winning is. It appears to them that it could be as high as 2/3 or as low as 0. Now consider the following pair of options:

	30 red balls	60 balls	
		black	yellow
Option V	\$100	\$0	\$100
Option W	\$0	\$100	\$100

In this example, most subjects prefer option W, because they 'know' that their chance of winning is 2/3, whereas their chance of winning with option V could be as low as 1/3 or as high as 1.

Note that subjects reversed their choice merely because the 'yellow' column was changed. According to the *independence principle*, you should ignore any column that has the same entries for both options. So your choice should not be affected by whether the 'yellow' column contains \$100 for both options or \$0. Hence, this pattern violates the independence principle.

Many people, nonetheless, feel a strong temptation to make the choices as Ellsberg's subjects (mostly economists) did, choosing X and W. Becker and Brownson (1964) have even found that subjects will pay money to avoid making choices in which the probabilities seem to be 'unknown.'

Kashima & Maher (1992; see also Maher, in press) examined a modification of the Ellsberg paradox in which you are first told whether or not the ball is yellow. Then, if the ball is not yellow, you have a choice between X and Y or between V and W. Ellsberg-type subjects presented with these modifications tended to chose X and V (thus not violating independence) rather than X and W. (Note that the switch cannot be explained in terms of information supplied by the fact that the ball is not yellow. If anything, that should raise the probability of a black ball and incline the subject toward W.) All that changed was the order in which information was revealed before the choice, yet the subjects could have anticipated such revelations at the time the original choice. Again, this fact suggests that ambiguity aversion is an error.

The Ellsberg example is a particularly clear case, but it is not isolated. Ambiguity enters many of our real decisions and opinions, such as those concerning the safety of nuclear power plants or the future of Soviet military policy. The ambiguity effects pits strong intuitions about an example against a powerful normative theory: that of expected-utility (EU) maximization. Many theorists (Shafer, 1976, 1981; Grdenfors & Sahlin, 1982) have taken it, or closely related phenomena, as a starting point for the development of what they take to be alternatives to EU theory and the Bayesian probability theory that it implies. Rawls (1971) argued for the worst-case decision rule in cases of ambiguity in the 'original position,' and the use of this rule provided a major argument for the difference principle, in which primary goods are distributed so as to benefit the least advantaged group.

This phenomenon and related examples demonstrating subjects' aversion to ambiguity has led to empirical research examining the causes and effects of ambiguity (Curley & Yates, 1985, 1989; Curley, Yates, & Abrams, 1986; Einhorn & Hogarth, 1985; Heath & Tversky, 1991), as well as theoretical work attempting to specify the relevance of ambiguity for normative models of decision making (reviewed by Camerer & Weber, 1992).

More generally, Ellsberg's (1961) seminal finding has been important because it calls into question three fundamental claims in utility theory, as presented by Savage (1954). Savage (1954) showed that the principle of maximizing EU could be derived from a set of seemingly uncontroversial axioms. Utility theory, as presented by Savage, consists of three related claims:

> Measurement Claim: subjective probabilities can be defined in terms of preferences among gambles;
>
> Descriptive Claim: utility theory describes people's behavior;
>
> Normative Claim: the rule of maximizing EU is a normative rule of decision making.

First, Savage showed that subjective probabilities could be defined in terms of preferences (if certain 'axioms' were true of sets of preferences). By defining probabilities in terms of preferences, Savage was able to develop the concept of subjective probability in a way that was acceptable to behaviorally oriented theoreticians.

Second, Savage proposed his theory as a descriptive model of choice under uncertainty. Utility theory was assumed to be a reasonably accurate model of people's choices under uncertainty. A crucial implication of this theory is that there is no meaningful distinction between 'known risk' and 'uncertainty.'

Finally, Savage showed that the principle of maximizing EU followed from a set of intuitively compelling axioms. Thus, Savage presented a strong justification of utility theory as a normative model. In Savage's (1954) theory, choice is a function of utilities and probabilities, where probabilities are one's subjective estimates of the likelihood of states of the world.

The ambiguity effect provides first-blush evidence against all three of Savage's claims. While previous discussions of ambiguity have noted the relevance of ambiguity for each of these claims, the implications have not been distinguished very clearly. In this paper, we shall discuss the implications of ambiguity for each of Savage's claims and show how the ambiguity effect leads to new insights into the uses and limits of utility theory. We conclude with a discussion of prescriptive implications, that is, implications for practice.

2 The measurement claim

There are two distinct ways of operationalizing the notion of 'degree of belief' or subjective probability (Ramsey, 1931). First, one can define a subjective probability as an intuitive judgment of probability. On this view, the way to measure subjective probabilities is to ask people. We can call this tradition the *introspective interpretation* of subjective probabilities. Alternatively, one can define subjective probability as a theoretical entity that is inferred from a person's choices. We can call this tradition the *behaviorist interpretation* of subjective probabilities, since probabilities are defined in terms of choices and are inferred from choices. One of Savage's contributions was to show that if certain constraints were true of people's preferences, then probabilities could be inferred from choices. Savage's theory was very useful to researchers wanting to apply the behaviorist interpretation.

When Ellsberg (1961) first discussed the issue of ambiguity, most researchers were committed to the behaviorist interpretation of subjective probabilities. Ramsey (1931) and others (Marschak, 1975) argued that one cannot just ask people for probability judgments. They claimed that if you ask people for probabilities, the answers you get are not necessarily meaningful. Both Ramsey (1931) and Savage (1954) suggested that people may not have access to intuitions about 'How likely is X.' They also suggested that such intuitions may have nothing to do with behavior. Ramsey argues: '..the quantitative aspects of beliefs as the basis of action are evidently more important than the intensities of belief-feelings.' (p. 171). Savage puts it quite clearly: 'Even if the concept were so completely intuitive, which might justify direct interrogation as a subject worthy of some psychological study, what could such interrogation have to do with the behavior of a person in the face of uncertainty, except of course for his verbal behavior under interrogation?' (p. 27).

Thus, when Ellsberg wrote his paper, many researchers believed that the only sensible way to define subjective probabilities was in terms of behavior. The ambiguity effect demonstrated that the probabilities inferred from choices are not coherent. That is, if a person states a pattern of preferences in which ambiguity is avoided (or preferred), then it is impossible to assign coherent probabilities to that person. For example, if the person prefers Gamble 1 whether black or white is associated with the payoff, then the probability of white in Gamble 2 must be less than the probability of white in Gamble 1, and the probability of black in Gamble 2 must be less than the probability of black in Gamble 1. If the probability of white and black in Gamble 1 are .5, then the probabilities of the two outcomes of Gamble 2 must add to less than one. Thus, Ellsberg's finding called into question the validity of the concept of subjective probability. This finding was obviously troubling to researchers in decision making. If observable behavior (choices) is the only type of admissible data and the probabilities inferred from this observable behavior are incoherent, then one could not develop a theory of decision making in terms of probabilities and utilities.

In the past two decades, researchers have become increasingly comfortable with the practice of asking people for probability judgments directly (e.g. Kahneman et al., 1982). Although such data were inadmissible 25 years ago (see e.g. Marschak, 1975), today the practice of asking subjects to give probability assessments is very common. For example, Kahneman and Tversky's work on judgment under uncertainty largely involves experiments in which subjects are explicitly asked for probability ratings. Perhaps this increased willingness to ask for numerical judgments is a result of the increase use of scoring rules both in theory (Lindley, this volume) and practice (Murphy & Winkler, 1977): scoring rules provide a behaviorist constraint on numerical judgments.

Nonetheless, we cannot assume that all problems with probability elicitation have been solved. Until they have been solved, and so long as hypothetical decisions are used to elicit probabilities, the ambiguity effect is relevant to probability measurement as well as to decision making.

3 The descriptive claim

Given that psychologists have become increasingly comfortable assessing probabilities directly as opposed to inferring them from preferences, it now makes sense to ask whether ambiguity affects preferences directly or through an effect on probabilities. Although this question would not have made sense to Savage or Ellsberg, it is a reasonable question today, since we accept as data both probability judgments and preference judgments.

An important goal of research on ambiguity is to explain why ambiguity influences choices in the ways it does. There have been two basic approaches to explaining ambiguity, one in which ambiguity affects beliefs (probabilities) and one in which ambiguity affects preferences directly.

Effects on belief. Some authors have attempted to explain ambiguity as an effect on belief. Einhorn and Hogarth (1985) account for ambiguity effects in terms of distortion of beliefs. In particular, when subjects are given an ambiguous probability - or some results that imply one - they use that probability as an anchor and they adjust it, as if they were adjusting toward some central point by regression. Adjustment is less when the anchor is near 0 or 1 than when it is more central. The central point itself depends on the subject and the situation; it may be taken as an index of optimism or pessimism when the outcomes differ in utility. The mechanism for this adjustment is the imagination of values both higher and lower than the anchor, and the averaging of these imagined values. The adjusted probability is then entered into the decision as if it were a stated probability.

In order to obtain results that support this view, subjects must not be told the overall 'marginal' probability, or else they must be discouraged from taking it too seriously. When this is done, most results support the theory. (Camerer and Weber, 1992, provide a thorough review.) The most important result is that the stated probabilities assigned to complementary events can systematically sum to less than one (Einhorn & Hogarth, 1985). For example, one subject, told that 4 witnesses has identified a car as blue and 1 had identified it as green gave .77 as the probability that it was blue. On another trial of the experiment, the subject gave .18 as the probability that the car was green, based on the same data (Table 4).

The Einhorn/Hogarth model predicts that adjustments resulting from ambiguity will be greater for more extreme probabilities. It therefore accounts for the fact that some people prefer to bet on an ambiguous urn when the probability of winning is very low: each urn contains 1000 balls; you win if #683 is drawn; in the unambiguous urn, the balls are numbered 1 to 1000; and in the ambiguous urn, each ball can have any number in that range. (According to Becker & Brownson, 1964, such preference for ambiguity was observed by Ellsberg; Einhorn & Hogarth, 1986, present additional supporting data.) According to the model, subjects assume that the probability of drawing #683 is greater in the ambiguous urn. Although this prediction has not been directly tested by asking subjects to compare the probabilities, no other model has been proposed to account for such findings.

The regression of belief strength toward some reference level is a reasonable strategy when evidence is poor. For example, if you are told that the probability of streptococcus infection is 10% in people with fever, sore throat, and swollen glands, but a recent study of 10 patients with sore eyes in addition found that 9 of them had this infection, a reasonable estimate of the true probability for a patient with all four symptoms would be closer to .40 than to .90. This regression heuristic can be overgeneralized to cases in which it is inappropriate, however. In the context of an experiment, adjusting beliefs is amounts to perversity when an experimenter specifies that they should not be adjusted. For example, when subjects are told that the probability of a disease depends on membership in a risk group, but membership in the risk group is unknowable and the overall probability of the disease is X (taking into account both members and nonmembers), the subject would be perverse not to accept the value of X as the probability. Of course, a subject who accepted the value of X might still *prefer* to bet on some other event with the same probability. That is a different issue.

In real life, however, probabilities are not so constrained. When the Food and Drug Administration tells us that the increased lifetime risk of cancer from some birth-control method is .003, we are *not* necessarily irrational to adjust that figure upward. In particular, we might have good reason to believe that such figures are generally underestimates and that the tendency to underestimate risk is greater when less information is available. Studies of changes in expert probability estimates as a function of increased data are needed. (Loewenstein & Mather, 1990, report such data for public perceptions.) It may well turn out that risk estimates first increase as a risk first enters our consciousness on the basis of preliminary findings and then decrease, in part because of regression to the mean (since studies would not be done if no risk is perceived) and in part because initial estimates are often based on 'worst case' assumptions. These assumptions, of course, are the result of psychological processes like those described by Einhorn and Hogarth. Thus, non-experts may be unwise to adjust reported probability estimates upward, if the estimates have already been revised upward once by the experts who produced them.

More generally, whether probability estimates should be adjusted depends on the social context in which they were generated. The rationality of adjustment depends on the facts of the matter in the social context. It could go either way. In sum, the adjustment of beliefs because information is ambiguous is a useful heuristic that may sometimes be overused.

Effects on preference. Although Einhorn and Hogarth (1985) present evidence demonstrating that ambiguity can influence choice through an effect on beliefs, this is not a sufficient explanation for all ambiguity effects. In Ellsberg-type experiments, at least the more sophisticated subjects can figure out the marginal probability for themselves, so accounts in terms of belief are unlikely to account for these results. Frisch (1988), Ritov and Baron (1990), and Heath and Tversky (1991) gave subjects the marginal probabilities, or asked subjects to provide them, so their results clearly demand an explanation in terms of preference rather than belief. (See Winkler, 1991, for other arguments concerning a preference account.)

Frisch and Baron (1988) provide an explanation for why ambiguity influences preferences, independent of beliefs. Ambiguity effects may be a result of our perception that important information is missing from the description of the decision (Frisch & Baron, 1988). Perhaps, then, we avoid ambiguous options because we really want to exercise another option: that of obtaining more information. (Roberts, 1963, p. 335, attributes this idea to Ward Edwards.) When this other option *is* available - as it often is - it is perfectly rational to choose it, providing that the information is worth obtaining. When the information is not available, however, or not worth the cost, we would do better to put aside our desire to obtain it and go ahead on the best evidence we have, even if it is 'ambiguous.' More generally, we can think of our tendency to avoid ambiguous decisions as a useful heuristic that points us toward the option of obtaining more information. From a *prescriptive* point of view, we probably do well to follow a rule of thumb that tells us to avoid irreversible commitments when information is missing. If we can learn to put this rule aside when the missing information is too costly or truly unavailable, however, we shall achieve our goals more fully in the long run.

Note that the effect of missing information is a matter of perception. In principle, an apparently unambiguous option could become ambiguous by calling attention to missing information. For example, in an urn with 50 red balls and 50 white ones, the probability of a red ball seem to be .5, without ambiguity. But think about the top layer of balls, from which the ball will actually be drawn. We have no idea what the proportion of red balls is in that layer; it could be anywhere from 100% to 0%, just like the proportion of black to yellow balls in the Ellsberg paradox. By thinking about the situation in this way, we have turned an unambiguous situation into an ambiguous one. The idea that some probabilities are 'objective' is simply a consequences of our not paying attention to unknown determinants of each event.

Support for our proposal comes from a study of hypothetical vaccination decisions (Ritov and Baron, 1990). In one experiment, subjects were told to imagine that their child had a 10 out of 10,000 chance of death from a flu epidemic, a vaccine could prevent the flu, but the vaccine itself could kill some number of children. Subjects were asked to indicate the maximum overall death rate for vaccinated children for which they would be willing to vaccinate their child. Most subjects answered well below 9 per 10,000. Of the subjects who showed this kind of reluctance, the mean tolerable risk was about 5 out of 10,000, that is, half the risk of the illness itself. The results are also found when the subject is asked to take the position of a policy maker deciding for large numbers of children. This result was interpreted as a biased toward omission, toward the default option of not vaccinating.

Of interest here is what happened when this manipulation was combined with ambiguity. In two experiments, subjects were told that the effect of vaccination, or of the flu, depended heavily on whether the child was in a 'risk group.' Children not in the risk group were save, but those in the risk group were subject to a considerable risk. The test for the risk group was not available. Thus, all that could be known was the overall probability of death in each case. The risk group was a form of salient missing information, which should, according to Frisch and Baron, induce a reluctance to choose the option in question. Subjects were in fact less willing to vaccinate when the result of vaccination was affected by membership in the risk group, thus supporting our hypothesis. Interestingly, the risk group did not affect preference when it applied to the effect of the flu. It seems that the effect of missing information reduces the tendency to act but has no effect on the tendency to omit action. This asymmetry deserves further investigation.

Heath and Tversky (1991) provide an account of ambiguity effects that is similar to that of Frisch and Baron (1988). They argue that people prefer to bet when their perceived competence is high. In several experiments, subjects were asked to give probabilities of answers to various questions, such as questions about general knowledge, football predictions, or political predictions. Subjects were then asked whether they would prefer to bet on their answers or on chance lotteries (based on colored poker chips) with the same probabilities. Subjects chose their answers when the probability they had assigned was high (indicating competence) or when they knew a lot about the subject. Subjects chose the lotteries when their probabilities were low or when they knew little. Heath and Tversky interpreted these results as follows: '... holding judged probability constant - people prefer to bet in a context where they consider themselves knowledgeable or competent than in a context where they feel ignorant or uninformed. We assume that our feeling of competence in a given context is determined by what we know relative to what can be known. Thus, it is enhanced by general knowledge, familiarity, and experience, and is diminished, for example, by calling attention to relevant information that is not available to the decision maker, especially if it is available to others' (p. 7).

They suggested that this competence effect has both cognitive and motivational determinants. Cognitively, the effect results from an overgeneralization of a rule that people do better in situations about which they have more information. Motivationally, Heath and Tversky suggest that the effect can result from anticipations of credit and blame: subject would expect more blame for a wrong guess on a lottery than for a wrong guess on an equally probable item in which the subject was expert. Such an expectation, however, requires the subject to assume that *others* are committing a cognitive error. Either this is true, in which case a cognitive error is being made somewhere, or not, in which case the subject is making an error in predicting the reactions of others. We also have no reason to think that subjects would expect others to make such an error, unless the error were often made. Thus, we regard the motivational account as secondary to some sort of cognitive account, if it is true.

Regardless of the source of this competence effect, its similarity to our earlier hypothesis is striking. In both accounts, the appearance of missing information leads to an unwillingness to bet, and Heath and Tversky's cognitive account is similar to our account in terms of overgeneralization of the reluctance to act when missing information might be available.

Some effects attributed to effects on beliefs (as postulated by Einhorn & Hogarth, 1985) might be at least partially the result of direct effects of perceived missing information on choice. For example, Kunreuther and Hogarth (1989) describe the effects of ambiguity on decisions about buying and selling (hypothetical) insurance contracts. In their experiments - done with actuaries as well as business-students - subjects set higher prices for insurance when risks were ambiguous. Ambiguity was manipulated by telling subjects that experts disagreed about the probability of the adverse event in question. Subjects were also told the mean of the experts' judgments, however, and this was held constant between ambiguous and unambiguous conditions. It therefore seems likely that subjects regarded the question of why the experts disagreed as missing information, so they were more reluctant to accept the risk.

In sum, two different mechanisms seem to produce ambiguity effects, one involving belief and the other involving preference. The former tends to moderate extreme beliefs when they are ambiguous. The latter inhibits people from choosing an option when they feel that information about its consequences is missing.

4 The normative claim

Savage (1954) provided a rationale for a normative theory which implies that uncertain states of the world are all assigned personal probabilities and decisions are consistent with the maximization of expected utility based on these probabilities. An important implication of Savage's theory is that '...for a 'rational' man - all uncertainties can be reduced to risks' (Ellsberg, 1961, p. 645). The ambiguity effect demonstrates that many people do make a distinction between different types of risk. Thus, people's intuitions are in conflict with a normative theory.

Central to Savage's theory is a form of the independence principle, which can itself be violated by people who are sensitive to ambiguity. Similar principles, along with the principle of transitivity, are used in later developments along the same lines (see Krantz et al., 1971, ch. 8).

Justification of the independence principle. Independence (in one form) requires an analysis of decisions into options, uncertain states of the world, and outcomes, which depend on the option and the state. According to the independence principle, if the option chosen does not affect the outcomes in some states of the world, then we can ignore the nature of these outcomes in those states. For example, in option A, you get a 1/1000 chance to win $1000 if a coin flip comes up heads, and $Z for sure if it comes up tails. In option B, you get $1 for sure if it comes up heads, and $Z for sure if it comes up tails. Z has the same value in both options. By the independence principle, you should make the same choice regardless of the value of Z, because in the state of the world 'tails' the outcome is the same regardless of my choice. Your choice really comes down to whether you prefer the dollar or the chance to win $1000.

More generally, the independence principle can be described in terms of a table like the following, in which the rows are the options and the columns are uncertain events or states of the world (as described by Jeffrey, 1983):

	state 1	state 2	state 3
option A	V	X	Z
option B	W	Y	Z

The entries in the table are the outcomes (V-Z). In the example just given, state 3 corresponds to tails, state 1 corresponds to heads and winning the lottery, state 2, to heads and losing. W and Y are both $1, V is $1000, and X is $0. By assumption, the entries in one column (state 3) are identical. The options therefore differ as a function of the choice only in the other columns. The independence principle states that the outcomes in the identical column (Z, here) should not affect the decision. The non-identical columns affect the decision in the same way, regardless of what is in the identical column.

If you follow independence and transitivity (plus other axioms that are less important), then you must make decisions as though you assigned probabilities to uncertain states of the world, assigned utilities to outcomes, multiplied the probability of each outcome by its utility, added up these products for the possible outcomes of each option, and chose the option with the highest sum (EU).

If you accept the axioms as constraints on your decision, then, normatively, you should not violate this EU formula.

How can the independence axiom be justified? One line of justification may be based on the definition of utility in terms of goal achievement (or, equivalently for this purpose, desire satisfaction). Importantly, we take utility to be a real property of states of the world, not an intervening variable designed to explain preferences. Thus, as Kahneman and Snell (1992) argue, judgments of utility are more like predictions than reports of inner states. When we make a judgment of the utility of an outcome, we are predicting how much that outcome will achieve all of our goals taken together. Note that, by this view, to say that two entries in the table are the same (e.g., to label them with the same letter) is to say that they are equivalent in terms of achieving goals.

Now, given this kind of table, we have two possibilities. Either the identical state (state 3) occurs or one of the non-identical states (state 1 or 2). If the identical state occurs, then the nature of the identical outcome (Z) does not affect the achievement of goals as a function of the option chosen, since the outcome is the same regardless of the option chosen. If one of the non-identical states occurs, then the nature of the identical outcome does not affect the achievement of goals either, because the identical outcome (Z) does not occur. Achievement of a goal is a matter of fact, so it depends on what is true of the world after the decision is made. (Recall that we have assumed that no goals concern counterfactual outcomes.) In sum, the nature of Z, the identical outcome, does not matter if Z occurs, and it does not matter if Z does not occur, so it does not matter. Independence therefore follows from the idea that rational decisions should be determined by the extent to which their outcomes achieve goals. (The same kind of argument can be used to defend related principles, such as those involving dominance or independence of irrelevant alternatives.)

Why people might still want to violate independence. The independence principle is usually illustrated with monetary outcomes, as in the Ellsberg paradox. When the entries in the table represent monetary outcomes, people may want to violate the principle for a couple of reasons. First, forgone or counterfactual outcomes affect their emotions, or more generally, the way in which consequences are experienced (before, during, or after they occur). For example, if Z is $1 in the table, then X ($0) could cause a feeling of regret, since you would realize that if you had chosen B you would have won something no matter what. If Z is $0, however, it will be easier for you to tell yourself that you might have won nothing anyway. Your experience of X is therefore changed by your knowledge of Z. In terms of goal achievement, then, X is no longer the same outcome for different values of Z. It should be represented with different symbols depending on the value of Z. Because the independence principle for goal achievement requires that X be the same consequence, the premise of the independence principle is not true, and you have not violated it if you make different choices for different values of Z. In sum, violations of independence (or of EU itself) that depend on emotional experiences need not be violations at all once the experiences are included in the descriptions of consequences. The trouble comes from describing the consequences as amounts of money. (Frisch & Jones, in press, make a similar point.) Evidence that subjects take such experiences into account in making decisions is summarized by Harless (1992).

In case it is difficult to imagine when the assumptions of the independence condition *are* met, consider the case of (what we shall call) Other decisions, in which each decision is made for another person, who does not know what the rejected options or counterfactual outcomes were, and in which we cannot assume that the Other has goals concerning the effect of these unknowns on choice (Baron, in press). If the decision maker truly took into account only the utilities of the recipient, not her own utilities connected with making the decision, the emotional effects of forgone or counterfactual outcomes would largely disappear. If the recipient's utilities concerned only the outcomes that he would know about, these effects would disappear completely.

For Self decisions (those made for the self), the fact that the same nominal outcome (e.g., '$1000) may lead to different real outcomes as a function of forgone options or counterfactual outcomes makes it difficult to test EU as a *descriptive* theory from behavior alone. To test the theory for Self decisions, we must measure the utility of outcomes *in the context of the decision itself*, using other methods than how people make decisions under risk (Baron, 1988, ch. 16). We can use the theory normatively and prescriptively in the same way, i.e., by describing the outcomes in the context of the whole decision and allowing its utility to depend on events that did not happen and options that were not chosen.

If we want the utilities of outcomes to be independent of the context, we do well to think about Other decisions. Our arguments in favor of the independence principle applied most clearly to this case. If we use Other decisions to test the theory descriptively, we will probably find all the same violations that have been found in Self decisions, such as the effect of certainty. Some experiments have used Other decisions decisions (Baron & Hershey, 1988; Kahneman & Tversky, 1984; Ritov & Baron, 1990; Spranca, Minsk, & Baron, 1991), finding that the theory still did not apply descriptively. In particular, the ambiguity effect found by Ritov and Baron (1990) was in the context of an Other decision, the vaccination of a child or (equally) a policy for vaccination of many children. When the conditions are met for the independence principle to apply, violations of that principle, such as the Ellsberg paradox, subvert the achievement of goals. In that sense, the pattern of choices observed in the Ellsberg paradox is nonnormative. We suggest that more research be done using Other decisions. Ambiguity effects in Self decisions are not clearly nonnormative. When these effects - and other effects - occur in Other decisions, they are more clearly nonnormative. They may be considered as overgeneralizations of heuristics that might be useful for Self decisions.

In sum, we have provided a defense of the independence principle in terms of goal achievement. This defense is intended as an answer to criticisms of the more traditional approach, which derives from the intuitive appeal of the axioms themselves (e.g., Slovic & Tversky, 1974).

5 The Allais paradox

The Allais paradox is another case in which the independence principle is violated (Allais, 1953). Consider the following gambles, in which the outcome is decided by drawing a ball at random from an urn containing 100 balls with the numbers 1 through 100 written on them:

	{Number on ball drawn}		
	1	2-11	12-100
{Situation X}			
Option 1	$1,000	$1,000	$1,000
Option 2	0	5,000	1,000
{Situation Y}			
Option 3	$1,000	$1,000	$0
Option 4	0	5,000	0

Many people in this situation are tempted to choose Option 1 in Situation X and Option 4 in Situation Y. In situation X, they are not willing to give up the *certainty* of winning $1,000 in option 1 for the chance of winning $5,000 in option 2: This extra possible gain would expose them to the *risk* of winning nothing at all. (If you do not happen to feel this way, try replacing the $5,000 with a lower figure, until you do. Then use that figure in choice 4 as well.) In situation Y, they reason that the difference between the two probabilities of winning is small, so they are willing to try for the larger amount.

This pattern of choices violates the independence principle. Balls 12-100 lead to the same outcome ($1,000) regardless of whether we choose Option 1 or 2 in Situation X, and they lead to the same outcome ($0) whether we choose Options 3 or 4 in Situation Y. By the independence principle, you should choose Options 1 and 3, or Options 2 and 4, but you should not choose Options 1 and 4. Usually, the independence principle is intuitively attractive, but many people are prone to violate it by choosing Options 1 and 4.

Shafer (1986) argues that it is not necessarily irrational to choose Options 1 and 4. He says that the 'constant' outcomes - those that are the same regardless of our choice - affect our goals or desires in the situation. (Lopes, 1987, makes a similar argument.) When we see that we can win a substantial sum of money for sure in Option 1, this reduces our desire for the larger sum. When we see that we are likely to lose no matter what, in Options 3 and 4, our desire to 'win big' increases.

This argument is less relevant if we change the example. Instead of the decision maker getting the money, it is donated anonymously and without explanation to his favorite nephew, or whoever. This is an Other decision. The nephew does not know what options were foregone or what states did not occur, so his experiences are unaffected by these things. Moreover (we assume), the decision maker has no reason to think that the nephew has any particular goals concerning options that were not chosen, or states that did not occur, in having decisions made on his behalf. The nephew's utilities thus cannot be affected by counterfactual outcomes, so Shafer's argument does not apply. We see here how Other decisions are, in a sense, simpler than Self decisions. By Shafer's account, the assumptions of the independence principle are typically not met in Self decisions, but it is easy to imagine how they might be met in Other decisions.

The perspective of Other decisions also strengthens another argument for the independence principle. Raiffa (cited in McClennan, 1983) points out that we may view the original problem as a sequential decision, as follows:

> First, a ball will be drawn out of an urn with balls numbered 1-100. If the number drawn is between 12 and 100 inclusive, the outcome is \$1,000 (for Situation X) or \$0 (for Situation Y). Otherwise, a second draw is made from a new urn with balls numbered 1 through 12. For Option 5, the outcome of the second draw is \$1,000, no matter what. For Option 6, the outcome is \$5,000 if the number is between 2 and 11 inclusive but \$0 if the ball is 1.

If we get to the stage of making the choice between Options 5 and 6, then the outcome for number 12-100 is irrelevant, for it did not occur. Raiffa argues that it should be irrelevant whether we make the decision before we know whether we get to the second stage of the game (as in the original Situations X and Y) or after we know (as in this example). McClennan (1983), points out that Raiffa and others who make similar arguments give no reason why the timing of the decision should not matter; they simply assert it, or suggest that most people's intuition would agree. But, to answer McClennan, it is clear that the timing would not matter to some Other who simply experienced the consequences without knowing the sequence of events that led to it (assuming that the Other has no goals concerning these non-experienced events).

6 Issues in application

We have argued that ambiguity effects can result from overgeneralization of heuristics concerning the postponement of decision making when information is perceived as missing. These effects can be nonnormative, that is, in opposition to the optimal achievement of our goals. But issues remain concerning the practical treatment of situations in which information *is* missing, for example, cases in which probability judgments disagree and we lack information about how to resolve the conflict. We discuss this problem here, as well as the problem of defining true probability in practical contexts, and the role of experts in decisions under ambiguity.

Conflict. Lindley, Tversky, and Brown (1979) have discussed the problem of conflicting judgments from a Bayesian point of view. Theoretically, they assume that judgments are a function of some underlying probability that we might call 'true.' If assumptions are made about the probability of each judgment given each possible true probability, then Bayes's theorem can be used to derive a probability distribution over the possible true probabilities, and the mean of this distribution can be taken as the best estimate. In this way, judgments made by different methods or by different people can be reconciled. Lindley et al. give several examples.

True probability. But the true probability is still known only probabilistically. In fact, the concept of true probability requires explication. The very distinction that inspired Ellsberg, that between uncertainty and risk (Knight, 1921; Luce and Raiffa, 1957) implies that some probabilities can be known with certainty but others cannot be known, only judged. This distinction lies at the heart of a number of recent alternatives to EU theory, reviewed by Camerer & Weber (1992). Bayesians in the tradition of Savage are skeptical about this distinction, however. They see cases of 'known risk' as merely convenient simplifications, in which various judges and methods agree closely on the probability. In some cases, this agreement has resulted from overwhelming of priors by extensive data. In other cases, it results from ignoring relevant data, as when a judgment is made about the probability of a certain patient having a disease on the basis of populations statistics, ignoring potentially relevant data about the individual. From this Bayesian point of view, the only possible 'true' probabilities are zero and one, and these apply mostly after the fact. Everything else involves judgments based on incomplete information.

This sort of Bayesian stance runs into conflict with our way of talking about probabilities. We say thinks like, 'I thought that the probability was X, but it was really Y.' In some cases, laws and regulations are stated in terms of probabilities, such as limits on the probability of disease caused by exposure to a chemical. These regulations are written as if the probability was an objective fact.

Brown (in press) proposed a Bayesian analysis of the idea of true probability, an analysis that allows such ways of talking to make sense. The true probability is the judgment that experts would converge on, as further relevant information became too costly to collect. In each case, the specific information required would differ, and a true probability need not exist in every case. For example, in determining the cancer risk from a chemical, the true probability might be thought of as the estimate derived from epidemiological data concerning cancer rate as a function of yearly exposure in the whole population of interest. Experts would form their priors on the basis of animal studies and theoretical beliefs about the form of the dose-response function. As more data were collected, these beliefs would approach the same asymptote. In principle, given sufficient time, data like these could be collected for different groups of patients. But such data on the interaction between exposure and individual characteristics would presumably be too costly to collect, so the population asymptote would be the one that experts would have in mind as the 'true probability.' Thus, Brown's analysis assumes both that an intermediate asymptote exists and that expert judgments would converge. He argues for the plausibility of these assumptions in many cases.

This definition of true probability avoids the conclusion that 'the true probability is always just 0 or 1.' It assumes that there is some sort of standard body of evidence that people want in each case. To take another example, a doctor might sensibly say, 'I can't assign a probability that the patient has cancer until we get back all the test results.' Here, the standard tests constitute the standard evidence. Note that a biopsy would be definitive here, but that is not included in the standard tests because it is considerably more costly (and perhaps because it would not make sense to speak of probabilities at all if it were available).

Brown's account fits neatly with our own theory of ambiguity as missing information. When a standard body of evidence exists and has not been obtained, people will be aware that this information is missing, and they will desire to collect it before acting. In most cases, this hesitation will be justified. In some cases, however, the situation will be classified as one in which the standard information is easily available, when, in fact, the information is not available at reasonable cost. From this perspective, then, ambiguity effects arise in situations seen as similar to those in which additional information is available. In the Ellsberg urn, for example, the proportion of balls is seen as something that is usually given. The unique aspect of the problem is that the experimenter won't tell.

Expert judgment vs. democracy. People fear risks that are not well known (Slovic et al., 1984). These risks include those of new technologies such as genetic engineering. Another example is the risks resulting from changes in legal standards: part of the U.S. 'liability crisis' of the 1980's was the unwillingness of insurance companies to write liability policies when court standards could change retroactively, as they did several times in recent years (Huber, 1988).

Hacking (1986) makes an argument with which many would probably sympathize. He is happy enough to have policy decisions made on his behalf by decision analysis when probabilities of relevant outcomes are well known, but not when probabilities are subjectively judged. Presumably, probabilities would be well known for things like the success rate of various medical therapies for various disorders. Probabilities would not be well known for events such as meltdowns of nuclear power plants (especially when their design is new). In cases of the latter type, we would have to rely more heavily on traditional methods of decision making, which stress participation of those affected, or holistic subjective judgments by elected representatives.

On the other hand, we have argued that missing information is always present whenever probabilities are involved. What changes from case to case is its psychological salience. Normatively, we ought to make decisions on the basis of our best estimate of the probability, it would seem.

An exception to this argument occurs when the risk to one person is correlated with the risk to another and when the utility function for harm is nonlinear with the number of people. Correlated risks are found in the case of disasters, e.g., hurricanes or earthquakes, since harm to one person from such a source implies that others are more likely to be harmed as well. But the argument as stated here applies to individuals.

As we have noted already, the 'best estimate' could be systematically biased against caution in the case of new technologies. Often, the best estimate is arrived at by trying to imagine all possible ways in which something could go wrong. Yet, as Fischhoff, Slovic, and Lichtenstein (1978) have shown, we might tend to err on the side of leaving things out because of our inability to think of them, and therefore estimate on the low side. The public's intuition that experts underestimate risks ('You've been wrong so many times before, so why should we believe you now¿) might be justified.

On the other hand, the public could be basing its judgment on a biased sample of cases that come to mind simply *because* the experts erred against caution, such as the Three-Mile-Island nuclear incident and the problems with some intrauterine devices. Perhaps as many, or more, cases could be found in which experts erred in favor of caution. Experts, too, could be sensitive to ambiguity effects. (The U.S. Food and Drug Administration is said to routinely boost risk estimates when the data on which they are based are in any way inadequate.)

In principle, these problems are remediable. Enough experience exists with risk estimates to allow a direct test of the existence of bias. Such tests have not been done. In the meantime, risk analysts ought to do the best they can. Perhaps they should correct for various sources of error. Putting this another way, our true best estimate should include a correction - if needed - for under- or overestimation as a function of the amount of information available. Analysts can also use risk analysis to determine when more data will be helpful and when it will not.

Political factors are sometimes relevant. One of the purposes of risk analysis is to help reduce political friction. For this purpose, the risk analysis ought to be open to criticism by the public. Conceivably, such criticism can improve the accuracy of risk analysis, but even if it impairs accuracy it might be worth soliciting. In addition to soliciting public input, risk analysis should also consider educating the public about such matters as the ambiguity effect just described. The intuition that 'we should not act until we know the probability' should be understood as one that has a legitimate basis only insofar as systematic bias enters the process of risk analysis or insofar as collection of additional data is worthwhile.

Intertwined with the ambiguity effect is also a bias toward the status quo, or toward inaction (Ritov & Baron, 1990). Ambiguity seems to exaggerate this bias (Ritov & Baron, 1990), but it is present in any case. The amount of money that people will pay to rid themselves of a risk they already have is far less than the amount that they will accept in order to take on the same risk (Thaler, 1980; Viscusi, Magat, & Huber, 1987). If people could learn to overcome this bias - and it seems that they can to some extent (Larrick, Morgan, & Nisbett, 1990) - we could take their resistance to new technology more seriously. The existence of this bias toward inaction therefore makes more plausible the claim that people subvert their own goals by favoring present risks over smaller risks that just happen to be new.

What of Hacking's argument? In cases in which the public has reason to distrust those in charge of a decision analysis, traditional methods of decision making might be better. As noted, self-serving bias - the basis of distrust - can be minimized by precautions surrounding the analysis itself. If substantial self-serving bias is absent, however, or if adequate precautions are taken to avoid its effects, perhaps Hacking would do well to trust his fate to the best guess of experts rather than to the political process. The political process itself is hardly perfect.

References

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axioms de l'cole amricaine. *Econometrica, 21*, 503–546.

Baron, J. (1988). *Thinking and deciding.* New York: Cambridge University Press.

Baron, J. (in press). *Morality and rational choice.* Dordrecht: Kluwer.

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology, 54*, 569–579.

Becker, S. W., & Brownson, F. O. (1964). What price ambiguity? Or the role of ambiguity in decision making. *Journal of Political Economy, 72*, 62–73.

Brown, R. V. (in press). Impersonal probability as an ideal assessment based on accessible evidence: a viable and practical construct? *Journal of Risk and Uncertainty.*

Camerer, C. & Weber, M. (1992). Recent developments in modeling preferences: uncertainty and ambiguity. *Journal of Risk and Uncertainty, 5*, 325–370.

Curley, S. P., & Yates, J. F. (1985). The center and range of the probability interval as factors affecting ambiguity preferences. *Organizational Behavior and Human Decision Processes, 36*, 272–287.

Curley, S. P., & Yates, J. F. (1989). Am empirical evaluation of descriptive models of ambiguity reactions in choice situations. *Journal of Mathematical Psychology, 33*, 397–427.

Curley, S. P., Yates, J. F., & Abrams, R. A. (1986). Psychological sources of ambiguity avoidance. *Organizational Behavior and Human Decision Processes, 38*, 230–256.

Einhorn, H. J., & Hogarth, R. M. (1985). Ambiguity and uncertainty in probabilistic inference. *Psychological Review, 92*, 433–461.

Einhorn, H. J., & Hogarth, R. M. (1986}. Decision making under ambiguity. *Journal of Business, 59*, S225-S250.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics, 75*, 643–699.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 330–334.

Frisch, D. E. (1988). *The effect of ambiguity on judgment and choice.* Doctoral dissertation, Department of Psychology, University of Pennsylvania.

Frisch, D., & Baron, J. (1988). Ambiguity and rationality. *Journal of Behavioral Decision Making, 1*, 149–157.

Frisch, D. & Jones, S. K. (in press). Assessing the accuracy of decisions. *Theory and Psychology.*

Grdenfors, P. & Sahlin, N.-E. (1982). Unreliable probabilities, risk taking, and decision making. *Synthese, 53*, 361–386.

Hacking, I. (1986). Culpable ignorance of interference effects. In D. MacLean (Ed.), *Values at risk* (pp. 136–154). Totowa, NJ: Rowman & Allanheld.

Hammond, P. H. (1988). Consequentialist foundations for expected utility. *Theory and decision, 25*, 25–78.

Harless, D. W. (1992). Actions versus prospects: the effect of problem representation on regret. *American Economic Review, 82*, 634–649.

Heath, C., & Tversky, A. (1991). Preference and belief: ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty, 4*, 5–28.

Huber, P. W. (1988). *Liability: The legal revolution and its consequences.* New York: Basic Books.

Jeffrey, R. C. (1965). *The logic of decision* (2nd revised edition). Chicago: University of Chicago Press.

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

Kahneman, D., & Snell, J. (1992). Predicting changing taste: do people know what they will like? *Journal of Behavioral Decision Making, 5*, 187–200.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*, 341–350.

Kashima, Y., & Maher, P. (1992). *Framing of decisions under ambiguity.* Manuscript, Department of Psychology, La Trobe University.

Knight, F. H. (1921). *Risks, uncertainty, and profit.* London: Macmillan.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. 1). New York: Academic Press.

Kunreuther, H., & Hogarth, R. M. (1989). Risk, ambiguity, and insurance. *Journal of Risk and Uncertainty, 2*, 5–35.

Larrick, R. P., Morgan, J. N., & Nisbett, R. E. (1990). Teaching the use of cost-benefit reasoning in everyday life. *Psychological Science, 1*. 362–370.

Lindley, D. V., Tversky, A., & Brown, R. V. (1979). On the reconciliation of probability assessments. *Journal of the Royal Statistical Association* A. *142*, 146–180 (with commentary).

Loewenstein, G., & Mather, J. (1990). Dynamic processes in risk perception. *Journal of Risk and Uncertainty, 3*, 155–175.

Lopes, L. L. (1987). Between hope and fear: The psychology of risk. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 20, pp. 255–295). New York: Academic Press.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions.* New York: Wiley.

Maher, P. (in press). *Betting on theories.* Cambridge: Cambridge University Press.

Marschak, J. (1975). Personal probabilities of probabilities. *Theory and Decision, 6*, 121–153.

McClennan, E. (1983) Sure thing doubts. In B. P. Stigum & F. Wenstp (Eds.), *Foundations of utility and risk theory with applications*, pp. 117–136. Dordrecht: Reidel.

Murphy, A. H., & Winkler, R. L. (1977). Can weather forecasters formulate reliable probability forecasts of precipitation and temperature? *National Weather Digest, 2*, 2–9.

Raiffa, H. (1961). Risk, ambiguity, and the Savage axioms: comment. *Quarterly Journal of Economics, 75*, 690–694.

Ramsey, F. P. (1931). Truth and probability. In R. B. Braithwaite (Ed.), *The foundations of mathematics and other logical essays by F. P. Ramsey.* New York: Harcourt, Brace.

Rawls, J. (1971). *A theory of justice.* Cambridge, MA: Harvard University Press.

Ritov, I., & Baron, J. (1990). Reluctance to vaccinate: omission bias and ambiguity. *Journal of Behavioral Decision Making, 3*, 263–277.

Roberts, H. V. (1963). Risk, ambiguity, and the Savage axioms: comment. *Quarterly Journal of Economics, 77*, 327–336.

Savage, L. J. (1954). *The foundations of statistics.* New York: Wiley.

Shafer, G. (1976). *A mathematical theory of evidence.* Princeton: Princeton University Press.

Shafer, G. (1981). Constructive probability. *Synthese, 48*, 1–60.

Shafer, G. (1986). Savage revisited. *Statistical Science, 1*, 463–501 (with discussion).

Slovic, P., Lichtenstein, S., & Fischhoff, B. (1984). Modeling the societal impact of fatal accidents. *Management Science, 30*, 464–474.

Slovic, P., & Tversky, A. (1974). Who accepts Savage's axioms? *Behavioral Science, 14*, 368–373.

Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology, 27*, 76–105.

Thaler, R. H. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization, 1*, 39–60.

Viscusi, W. K., Magat, W. A., & Huber, J. (1987). An investigation of the rationality of consumer valuation of multiple health risks. *Rand Journal of Economics, 18*, 465–479.

Winkler, R. L. (1991). Ambiguity, probability, preference, and decision analysis. *Journal of Risk and Uncertainty, 4*, 285–297.